HEALTH EFFECTS INSTITUTE

Includes Errata Sheet Of 11-01-01

Part II: Sensitivity Analyses

Daniel Krewski, Richard T Burnett, Mark S Goldberg, Kristin Hoover, Jack Siemiatycki, Michael Jerrett, Michal Abrahamowicz, Warren H White, and Others

Continued

∞ *Errata* ∞

Created 11/01/01

Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality

A Special Report of the Institute's Particle Epidemiology Reanalysis Project

Final version, July 2000
Posted to the HEI Website 10/27/00

Page 161. Part II. Caption for Figure 5 should read:
City-specific relative risks in the ACS Study.

Page 162. Part II. Caption for Figure 6 should read:
Shape of concentration-response function (with standardized residuals plotted) for cities in the ACS Study.

Page 174. Part II. Table 32. After "O_3 (ppb)" in the left column, append footnote [b] that reads:
"[b] Based on daily 1-hour maximum concentrations."

Page 178. Part II. Table 33. For O_3 (second row from bottom), in the column "Description of Covariate and Source of Data", the entry should read exactly like the other three:
"Daily average concentrations averaged by year for 1980; from residential, commercial, or mobile monitors"

Page 259. Health Review Committee's Commentary. ***Gaseous Copollutants*** section. The third sentence should read:
"For four gaseous copollutants (carbon monoxide, nitrogen dioxide, ozone, and sulfur dioxide), city-specific annual means of daily average concentrations from the year 1980 were obtained from AIRS and used in the reanalysis (see Appendix E, Part II)."

At the end of the same paragraph, add this sentence:
"For this analysis, the ozone values were based on daily 1-hour maximum concentrations."

Part II, Appendix E (available on request)
Page 5. ***Gaseous Copollutants*** section. The second sentence should read:
"Daily average concentrations of NO_2, sulfur dioxide, ozone, and carbon monoxide were obtained from 1980 to 1989, in addition to the daily one-hour maximum concentrations of ozone."

Part II: Sensitivity Analyses

Daniel Krewski, Richard T Burnett, Mark S Goldberg, Kristin Hoover, Jack Siemiatycki, Michael Jerrett, Michal Abrahamowicz, Warren H White, and Others

THE HARVARD SIX CITIES STUDY

QUALITY ASSURANCE AUDIT OF THE DATA FOR THE HARVARD SIX CITIES STUDY

An independent Audit Team (led by Ms Kristin Hoover; see Appendix A to Part I) conducted a detailed audit of all data used in the analyses reported by the Original Investigators (Dockery et al 1993*; referred to as the Part I data quality audit), in addition to auditing the new variables used in the Reanalysis Team's sensitivity analyses. We designed the Part I data quality audit to provide an overview of the databases and an assessment of the data management procedures used by the Original Investigators. The Part I audit also assessed the accuracy of data in the analytic files used in the original analyses relative to the original data from which they had been derived. Our objective in the Part II data quality audit was to evaluate the accuracy of the new variables selected by the Reanalysis Team for inclusion in its sensitivity analyses. For both Parts I and II, we randomly selected 250 subjects whose questionnaires became the basis of the data quality audit. Part I included an additional random sample of 250 death certificates; these were used to audit the nosologic coding of each underlying cause and date of death. We selected a sample size of 250 in order to provide reasonable statistical accuracy for achieving the goals of the data quality audit. Specifically, we selected this sample size to provide almost complete certainty of finding an error as small as 1% (Y Wang et al, unpublished data, 1995), to distinguish between error rates of 1% and 5% with reasonable confidence, and to estimate error rates within about two percentage points of the true value. (Further details are provided in Appendix A of Part I.)

For the Part II data quality audit, we included 17 variables from the initial questionnaires, 5 variables from follow-up questionnaires completed at 3, 6, and 12 years after enrollment into the study (these were not used in the original paper), and 2 variables derived from measurements of pulmonary function conducted at the time of enrollment. In addition, for the 60 subjects selected for the questionnaire audit and who had died during the follow-up period, we audited the underlying cause of death from death certificates obtained by the Original Investigators. The audit also examined the time of subjects' first move outside the original city of residence, on the basis of residence histories that the Reanalysis Team coded; we used these data in our assessment of population mobility in the Six Cities Study.

Part II Audit

We audited variables for the Part II analysis by comparing selected variables from the initial questionnaire that had been completed at the time of enrollment, as well as some other selected variables from the follow-up questionnaires, to the data in the electronic analysis file provided to the Reanalysis Team. We evaluated underlying causes of death using death certificates obtained by the Original Investigators for 60 subjects known to have died out of the 250 subjects in the random sample of audited questionnaires. We found no errors in variables for bronchial asthma, city of residence, date of birth, amount of wine/liquor consumed, marital status, race, or underlying cause of death. Variables in which we detected errors include occupation code from census, industry code, number of years living in same town, chest illness, alcohol consumption (multiple variables), age started smoking, number of packs of cigarettes smoked per week, number of years of smoking cigarettes, and heart trouble or high blood pressure.

Table 1 summarizes the variables in error (in alphabetical order by SAS variable name [SAS Institute, Cary NC] from the analysis file), and includes comments about these errors. (A more detailed presentation is in Appendix A,

* The original article appears in its entirety at the end of this Special Report.

This is one section of an HEI Special Report that includes an HEI Statement about the research project, a Preface to the Particle Epidemiology Reanalysis Project, the Investigators' Report (Introduction, Summary, Part I, and Part II), a Commentary by the Institute's Health Review Committee, and the Original Publications and Comments on the Reanalysis by the Original Investigators. Correspondence concerning *Part II: Sensitivity Analyses* may be addressed to Dr Daniel Krewski, Professor of Epidemiology & Statistics, Department of Epidemiology & Community Medicine, Room 3229C, 451 Smyth Road, University of Ottawa, Ottawa Ontario K1H 8M5, Canada.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award R824835 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

Table 1. Findings from the Phase II Audit of the Initial Study Questionnaires[a] from the Six Cities Study

SAS Variable Name from the Analysis File	Description of Variable	Number (and %) of Errors Found in 249 Questionnaires	Number (and %) of Errors Found in 89 Questionnaires by Original Investigators' Internal Audit (1981)	Type of Error Noted in Phase II Audit
AGECIG	Age started smoking: 0 = nonsmokers; ages 1–75 allowed by coding	1 (0.4)	0 (0.0)	Apparent coding error
BEER	Beer: 0 = none; 1 = < 200 oz/wk; 2 = > 200 oz/wk	2 (0.8)	1 (1.1)	Apparent coding errors
CHSTIL1	Chest illness diagnosed by doctor: 0 = no for bronchitis, emphysema, or pneumonia; 1 = yes for bronchitis; 2 = yes for emphysema; 4 = yes for pneumonia; higher numbers for subjects diagnosed with two or more diseases	4 (1.6)	3 (3.4) HSPH's audit concluded that error rate for this variable had not resulted from any systematic problem, so no recoding had been done.	Apparent coding errors
CIGWK	Number of packs of cigarettes smoked per week (20 cigarettes/pack)	3 (1.2)	3 (3.4)	Apparent coding errors
DRINK	Present use of alcoholic beverages: 0 = no; 1 = yes; part B asks if use is as often as 1 day/wk, for which 0 = no, 1 = yes, 2 = sum of both yes scores	1 (0.4)	0 (0.0)	Apparent coding error
HBP	Heart/blood pressure trouble: Has doctor ever diagnosed high blood pressure or heart problems? If yes, has this been treated in the last ten years? Scores could total as high as 8	4 (1.6)	0 (0.0)	Apparent coding errors
IND	Industry code	5 (2.0)	11 (12.4) HSPH's audit stated that retired, disabled, and unemployed subjects could not be distinguished, which resulted in many errors in interpretation. Other common errors: Working wives were often coded as housewives without reference to outside employment; unjustified assumptions were made about jobs when no information was available as to specific duties. Documents show efforts to correct errors.	Discussed in detail in Appendix A[b]
OCC	Occupation code (documents show that this variable was later superceded by another code)	5 (2.0)	21 (23.6) Documents show efforts to correct errors.	Discussed in detail in Appendix A
YRSCIG	Total years smoked cigarettes	2 (0.8)	0 (0.0)	Apparent coding errors
YRSHERE1	Number of years resident in this town	5 (2.0)	7 (7.9) Audit noted that consistent coding rules had not been carefully followed, and that years in military service should have been subtracted. Years in same city were counted even if not continuous.	Discussed in detail in Appendix A

[a] A total of 249 baseline questionnaires were available to audit the variables listed in this table. In addition, the Audit Team was able to extract information from follow-up questionnaires to confirm variables for marital status, race, city of residence, and date of birth.

[b] Appendix A is available on request from the Health Effects Institute.

which is available on request from the Health Effects Institute.) We audited five variables not included in the Original Investigators' published paper from follow-up questionnaires that had been completed 3 and 6 years after enrollment. These variables included height, weight, smoking history, number of years of cigarette smoking, and number of packs of cigarettes smoked per week. Audit of the analysis file for the height (HT) variable from the 3-year follow-up questionnaire revealed three errors in 249 questionnaires examined (1.2% error rate). For two subjects the height data for years 3 and 6 had been switched, which also caused an error in year 6 (0.8% error rate). The third had an incorrect entry for the year 3 questionnaire. One rounding error was noted in year 6 data when we audited the weight (WT) variable for the 3- and 6-year follow-up intervals, producing an error rate for year 6 of 0.4% (1/250). We observed no errors at the 3-year follow-up interval for any of the smoking variables (smoking status [SMOK], number of packs of cigarettes smoked per week [CIGWK], and number of years of cigarette smoking [YRSCIG]). There were no errors in SMOK at the 6-year follow up. We noted one rounding error in year 6 for YRSCIG, which resulted in an error rate of 0.4% (1/250), and there was one incorrect entry (0.4%; 1/250) in CIGWK at the 6-year follow up.

We audited three variables (HT, WT, CIGWK) from the last follow-up questionnaire, which had been completed 12 years after subjects had been enrolled in the study. A total of 247 questionnaires were available for year 12 (3 missing); we observed no errors in any of the variables with the possible exception of one case in which the entries for height and weight appeared to have been reversed on the questionnaire.

Summary of Audit Findings

The Audit Team found no errors in these data that would induce important effects in the statistical analyses (ie, errors in excess of 5%; the highest error rate was 2.4%). Coding of residential histories was done by subcontractors to the Reanalysis Team; the error rate in the coded variables for the date subjects first moved outside the original city of residence was 3.6%. Five of the nine observed discrepancies involved an error of 1 calendar year in the date of the first move. Although this error rate was somewhat higher than those for the original studies, it was still less than 5%. We thus concluded that the data were of sufficient quality for the purposes of the Part II sensitivity analyses.

ALTERNATIVE RISK MODELS

The Six Cities Study Original Investigators' Analytic Approach

Using Cox proportional-hazards regression models of survival, the Original Investigators (Dockery et al 1993) had examined the association between mortality in the Six Cities Study cohort and ambient air quality, as indexed by fine particles ($PM_{2.5}$)*, sulfate (SO_4^{2-}), total suspended particles (TSP), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), ozone (O_3), and aerosol acidity (H^+). Positive associations were observed with all measures of air pollution except ozone, and fine particles displayed the strongest association with mortality of all the measures examined; consequently, the Original Investigators had focused their analysis on this pollutant. In our reanalysis, we also focused on this pollutant in order to examine the robustness of this association when specifying models with different determinants of mortality and when applying different statistical approaches.

An assumption of the Six Cities Study Original Investigators' survival model had been that the relative increase in the underlying hazard function, or instantaneous rate of death, was constant over the entire follow-up period and was modulated by a number of risk factors for mortality such as smoking habits, education, and air pollution. The time axis for this survival analysis had been calendar year (1974 through 1989).

Effects of gender and age at enrollment in the study had been accounted for in the analysis by stratifying the baseline hazard function according to different categories of the covariates; age had been stratified on the basis of 5-year age groups. Because over 95% of the cohort was white, only whites had been included in the original analysis. The mortality risk factors that had been considered in the Original Model used by the Original Investigators of the Six Cities Study are listed in Table 2.

In addition to overall mortality, the mortality rate ratios had also been examined by the Original Investigators for the following underlying causes as defined in the *International Classification of Diseases, Ninth Edition* (ICD-9; World Health Organization 1975): cardiopulmonary diseases (ICD-9 codes 400–440, 485–496), lung cancer (ICD-9 code 162), and all other causes excluding cardiopulmonary disease and lung cancer. The Original Investigators used "mortality rate ratios" (Dockery et al 1993) and "mortality risk ratios" (Pope et al 1995) to describe the association between air pollution and mortality. Both

* A list of abbreviations and other terms appears at the end of the Investigators' Report.

Table 2. Covariates Included in the Original, Full, and Extended Models for the Reanalysis of the Six Cities Study[a]

Covariate	Alternative Risk Model		
	Original	Full	Extended
Tobacco consumption			
Current-smoker[b]	✓	✓	✓
Current-smoker years of smoking		✓	✓
(Current-smoker years of smoking)^2		✓	
Current-smoker cigarettes per day		✓	✓
(Current-smoker cigarettes per day)^2		✓	
Current-smoker pack-years	✓		
Former-smoker[b]	✓	✓	✓
Former-smoker pack-years	✓	✓	✓
(Former-smoker pack-years)^2		✓	
Age started smoking (current-smokers) ≤ 18 years[b]		✓	✓
Age started smoking (current-smokers) > 18 years[b]		✓	✓
Education level			
High school versus less than high school[b]		✓	✓
More than high school versus less than high school[b]		✓	✓
Less than high school versus high school or more than high school[b]	✓		
Exposure to dust or fumes[b]	✓	✓	✓
Body mass index	✓	✓	✓
(Body mass index)^2		✓	✓
Marital status			
Married versus single[b]		✓	✓
Separated versus single[b]		✓	✓
Widowed versus single[b]		✓	✓
Alcohol consumption			
Beer consumption [b]		✓	✓
Wine consumption[b]		✓	✓
Liquor consumption[b]		✓	✓
Interaction with gender			
Current-smoker[b]		✓	
Current-smoker years of smoking		✓	
(Current-smoker years of smoking)^2		✓	
Current-smoker cigarettes per day		✓	
(Current-smoker cigarettes per day)^2		✓	
Current-smoker pack-years	✓		
Former-smoker[b]		✓	
Former-smoker pack-years		✓	
(Former-smoker pack-years)^2		✓	
Age started smoking (current-smokers) ≤ 18 years[b]		✓	
Age started smoking (current-smokers) > 18 years[b]		✓	
High school versus less than high school[b]		✓	
More than high school versus less than high school[b]		✓	

(Table continues next page)

[a] All three of these models were analyzed with standard Cox proportional-hazards regressions.

[b] Dichotomous (yes/no) variable.

Table 2 *(continued)*. Covariates Included in the Original, Full, and Extended Models for the Reanalysis of the Six Cities Study[a]

Covariate	Alternative Risk Model		
	Original	Full	Extended
Interaction with gender (*continued*)			
Occupational exposure to dust or fumes[b]			
Body mass index		✓	
(Body mass index)^2		✓	
Married versus single[b]		✓	✓
Separated versus single[b]		✓	✓
Widowed versus single[b]		✓	✓
Beer consumption[b]		✓	✓
Wine consumption[b]		✓	✓
Liquor consumption[b]		✓	✓

[a] All three of these models were analyzed with standard Cox proportional-hazards regressions.

[b] Dichotomous (yes/no) variable.

terms refer to the ratio of the mortality rate at a higher level of air pollution relative to the mortality rate at some lower level. (Under the proportional hazards assumption made by the Original Investigators, this ratio is constant over time.) The Original Investigators found it convenient to use the pollution levels in the cities with the highest and lowest ambient air pollution levels as the basis for calculating the ratio of mortality rates. Unless otherwise specified, we follow this practice and use the term relative risk to denote the mortality risk ratio.

Note the relative risk can be calculated using the data from only two cities with the highest and lowest pollution levels, or by fitting an exposure-response model to the data for all cities together, and then evaluating the relative risk at the average pollution levels observed in the most-polluted and least-polluted cities. In most cases, relative risks reported by the Reanalysis Team are based on fitted exposure-response models.

Estimates of the log–relative risks had been obtained by maximizing the partial likelihood function of the Cox proportional-hazards model. Confidence intervals (95%) for the log–relative risks had been calculated under the assumption that they were normally distributed; that is, by adding and subtracting 1.96 times the standard error of the estimated regression coefficient.

The Reanalysis Team's Analytic Approach

The Reanalysis Team considered a number of alternative risk models that included additional covariates not examined in the original analysis; we also considered different functional forms or categorizations of original covariates, and 1-year age groups to stratify the baseline hazard function.

In our reanalysis, the Team also used age as the time axis, with age at enrollment into the study and age at event (death or censoring) modeled with respect to air pollution and other determinants of mortality. This approach has been shown to more fully capture the effects of age on survival than does using calendar year as the time axis (Breslow and Day 1987).

The Reanalysis Team initially considered a Base Model (with stratification by age and gender) that included air pollution with no additional determinants of mortality. We also included several additional covariates in a new regression model (the Full Model, Table 2). The Team included quadratic terms of a number of continuous variables that might have nonlinear effects, such as number of packs of cigarettes smoked, years of smoking, and body mass index (BMI); we also included other variables, not considered by the Original Investigators, that accounted for age at which smoking started and marital status. Because we wished to examine the effects of educational attainment in more detail, we considered three levels of attained education (less than high school, high school, and more than high school). The Team took into account the possibility that the effects of these risk factors could vary by gender by including an interaction term for each of these factors.

We then developed a more parsimonious model by removing those variables that did not significantly improve the goodness of fit. In particular, we dropped any

covariate from the Full Model if the *P* value derived from an increase in the log-likelihood function when we removed the covariate was greater than 0.05 (ie, likelihood ratio test). We continued this procedure until there was no further statistical justification for removing any other covariate. Regardless of the results of the likelihood ratio test, we retained a covariate when the corresponding gender interaction was statistically significant (Wald test $P < 0.05$). The parsimonious model derived in this way for all-cause mortality is referred to as the Extended Model. The Team also used this set of covariates to model mortality for cardiovascular disease (ICD-9 codes 400–459), respiratory disease (ICD-9 codes 460–519), lung cancer (ICD-9 code 162), other types of cancer excluding lung (ICD-9 codes 140–161, 163–239), and all remaining causes.

We also examined indicators of pulmonary function, forced vital capacity (FVC) and forced expiratory volume in one second (FEV_1), that the Original Investigators had obtained but had not included in their original analysis. We considered only pulmonary function data obtained at the time of enrollment because follow-up tests that had been conducted during the course of the study were judged to be a new analysis and thus outside the terms of reference of the reanalysis. We incorporated these variables by first carrying out a regression of the natural logarithm against the logarithms of height and age, thereby obtaining predicted pulmonary function values specific to the height and age of each individual in the study. We then included the residuals (observed minus predicted logarithmic pulmonary function volumes) from these models as determinants of mortality in the Cox proportional-hazards regression models.

Testing the Cox Proportional-Hazards Assumption

The validity of the Cox proportional-hazards assumption was evaluated in all models using test statistics provided in the statistical computing software S-PLUS (Grambsch and Therneau 1994). This test examines departures from the Cox proportional-hazards assumption in a linear manner. (Nonlinear departures from proportionality are examined in the Flexible Modeling section.) Although we found no statistical evidence of departures from the Cox proportional-hazards assumption in any model we examined ($P > 0.2$) using either calendar year or age as the time axis, the relative risk of mortality for fine particles varied slightly from a linear association that is consistent with the assumption of proportional hazards with both calendar year and age (Figure 1).

Relative risks of mortality associated with an increase in ambient fine particles are shown in Table 3 according to model specification (Base, Original, Full, and Extended), time axis used in the Cox model (calendar year or age), and cause of death (all causes, cardiopulmonary disease, cardiovascular disease, respiratory disease, lung cancer, other cancers, and other causes). The relative risks provided in Table 3 were scaled to estimate relative risks across the range of distribution levels of $PM_{2.5}$ (18.6 $\mu g/m^3$), the benchmark used by the Original Investigators.

Adjusting covariates using either time axis (age or calendar year) reduced the relative risk for each underlying cause of death, except for other cancers, for which a small increase was observed using the Full and Extended Models. We found that the relative risks in all three alternative risk models (Original, Full, and Extended) were similar.

The Reanalysis Team found that relative risk of mortality associated with an increase in fine particles had the following ranking among the underlying causes of death: lung cancer > cardiovascular disease > cardiopulmonary disease > all causes > other causes > other cancers > respiratory disease. Formal statistical significance ($P < 0.05$) was achieved for all causes and for cardiovascular and cardiopulmonary disease, in part because of the greater number of deaths in these categories than in other disease groupings. [The relative risk associated with fine particles was slightly higher if the underlying cause of death was restricted to ischemic heart disease (ICD-9 codes 410–414), with relative risk of 1.43 (95% CI: 1.06–1.92), based on the Extended Model and calendar year as the time axis (data not shown).] This result suggests that particulate air pollution may be affecting people with heart diseases more than it affects those with vascular problems.

The Reanalysis Team examined the effect of health status at enrollment on the association between mortality and fine particle air pollution by including adjusted FVC or FEV_1 as a covariate in the Extended Model using calendar year as the time axis for all causes of death. Both FVC and FEV_1 were strong predictors of mortality. A reduction in FVC corresponding to a change in the ratio of FVC to its adjusted value from 1 to 0.85 (representing a clinically significant reduction) resulted in a relative risk of death of 1.33 (95% CI: 1.28–1.39). The corresponding relative risk of a similar decrease in FEV_1 was 1.22 (95% CI: 1.18–1.25). However, the effect of fine particles on mortality was not appreciably altered by adjustment for FEV_1; RR = 1.27 (95% CI: 1.09–1.49) as compared with RR = 1.26 (95% CI: 1.08–1.47) prior to adjustment. Adjustment for FVC also did not influence the effect of fine particles on mortality (RR = 1.19, 95% CI: 1.11–1.52).

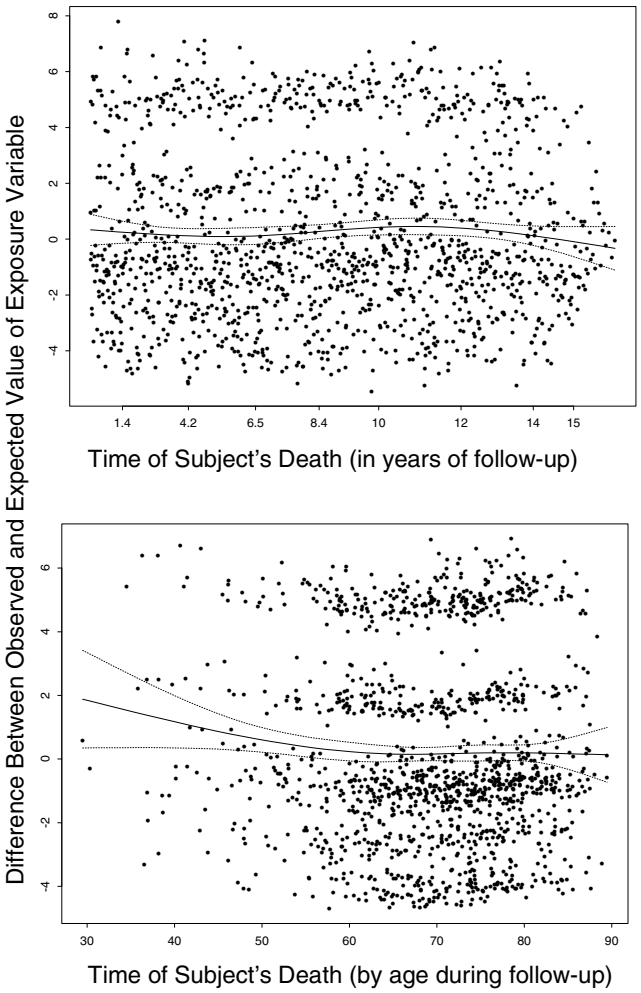

Figure 1. Proportional-hazards model assumptions for two time axes in the Six Cities Study. Log–relative risks due to each failure time (or time of death) [$\beta(t)$] for fine particles are plotted; relative risk estimates are based on the Extended Model. The y axis in both panels represents the difference between the observed value of the exposure variable for the person who died and the value expected on the basis of the fitted model. Panel A: Time of subject's death on the basis of years of follow-up. Panel B: Time of subject's death on the basis of age. Spline function smoothing of the association between log–relative risk and the time axis is shown by the solid line; the 95% confidence interval is shown by the dashed lines.

IDENTIFICATION OF SENSITIVE SUBGROUPS

Ambient air pollution, as indexed by fine particles, was associated positively with mortality from all underlying causes of death. To explore this finding in greater depth, the Reanalysis Team examined the association between particles and mortality within a number of cohort subgroups in order to identify those that may be more or less susceptible to the effects of ambient air pollution.

The relative risks for all-cause mortality associated with an increase in $PM_{2.5}$ of 18.6 $\mu g/m^3$ are shown in Table 4 for selected personal characteristics. We derived these estimates using the Extended Model with calendar year as the time axis and stratifying the baseline hazard function by 1-year age groups and gender. The relative risk of death associated with exposure to fine particles decreased with educational attainment and age; and it was higher in those people who reported workplace exposure to dust or fumes, less for married persons, greater for males than for females, greater for those subjects with self-reported heart or lung disease at time of enrollment, and greater for those individuals with compromised lung function. However, none of these interactions with air pollution achieved statistical significance ($P > 0.2$, likelihood ratio test). Fine particle association with mortality was insensitive to smoking status.

The Reanalysis Team also examined the influence of each of the six cities on the relative risk from fine particles by individually excluding each city from the analysis (see Table 4). The relative risks varied little after exclusion of any single community, with a range of 1.26 (excluding Portage) to 1.31 (excluding Steubenville). We note, however, that the 95% confidence interval (CI) in the relative risk included unity when Steubenville was omitted from the analysis. The associated relative risk of 1.31 was the highest among all cities in this influence analysis, indicating that the residents of Steubenville were dying at a lower rate than would be predicted by their air pollution exposure. However, exclusion of Steubenville also reduced the range in city-specific average $PM_{2.5}$ levels from 18.6 $\mu g/m^3$ to 9.8 $\mu g/m^3$, thereby increasing the standard error of the log–relative risk estimate and in turn widening the confidence interval.

Because the attained level of education appeared to have the strongest effect on the fine particle–mortality association, we examined the modifying effect of education in relation to the effect of other personal characteristics. Specifically, Table 5 shows the relative risk of all-cause mortality associated with increases in $PM_{2.5}$ of 18.6 $\mu g/m^3$, stratified on selected personal characteristics and educational attainment (high school or less, more than high school). These estimates are adjusted for all covariates included in the Extended Model.

The relative risk of mortality associated with fine particles was greater among individuals with high school education or less, compared to those with more than high school education in all subgroups examined except the "Other" marital status group; the relatively few subjects (517) in this group led to unstable estimates of risk (95% CI: 0.63–5.61). In the case of subjects under 40 years of age with more than high school education, the relative risk was higher (3.80) th9an for subjects with less education (1.54); however, there was considerable uncertainty in

Table 3. Relative Risks of Mortality by Cause of Death Associated with an Increase in Fine Particles in Risk Models with Alternative Time Axes in the Reanalysis of the Six Cities Study[a]

	Time Axis	
Alternative Risk Model[b]	Calendar Year	Age
All Causes [100%]		
Base	1.33 (1.14–1.54)	1.33 (1.15–1.55)
Original	1.29 (1.11–1.50)	1.29 (1.11–1.50)
Full	1.27 (1.09–1.49)	1.27 (1.09–1.48)
Extended	1.28 (1.09–1.49)	1.27 (1.09–1.48)
Cardiopulmonary Disease [54%]		
Base	1.39 (1.13–1.70)	1.39 (1.14–1.71)
Original	1.35 (1.10–1.66)	1.34 (1.09–1.65)
Full	1.31 (1.06–1.62)	1.30 (1.05–1.60)
Extended	1.32 (1.07–1.63)	1.31 (1.06–1.61)
Cardiovascular Disease [47%]		
Base	1.43 (1.15–1.78)	1.44 (1.16–1.79)
Original	1.41 (1.13–1.76)	1.40 (1.12–1.74)
Full	1.38 (1.10–1.72)	1.35 (1.08–1.69)
Extended	1.39 (1.11–1.73)	1.37 (1.09–1.70)
Respiratory Disease [7%]		
Base	1.11 (0.62–1.97)	1.10 (0.63–1.95)
Original	0.93 (0.51–1.71)	0.95 (0.53–1.72)
Full	0.89 (0.47–1.67)	0.94 (0.51–1.73)
Extended	0.88 (0.47–1.64)	0.93 (0.51–1.69)
Lung Cancer [8%]		
Base	1.53 (0.91–2.55)	1.64 (0.99–2.72)
Original	1.31 (0.76–2.25)	1.53 (0.90–2.60)
Full	1.30 (0.76–2.23)[c]	1.42 (0.84–2.42)
Extended	1.29 (0.75–2.22)[c]	1.45 (0.85–2.47)
Other Cancers [20%]		
Base	1.05 (0.74–1.48)	1.04 (0.73–1.47)
Original	1.04 (0.73–1.47)	1.02 (0.72–1.45)
Full	1.11 (0.78–1.59)	1.09 (0.77–1.55)
Extended	1.10 (0.77–1.57)	1.08 (0.76–1.54)
Other Causes [18%]		
Base	1.19 (0.80–1.75)	1.15 (0.78–1.70)
Original	1.16 (0.79–1.72)	1.12 (0.76–1.65)
Full	1.16 (0.78–1.73)	1.10 (0.74–1.63)
Extended	1.15 (0.77–1.71)	1.10 (0.74–1.62)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the Six Cities Study, this difference for fine particles was 18.6 $\mu g/m^3$. Causes of death are shown with percentage of all causes. Data are RRs with 95% CIs.

[b] See the Alternative Risk Models section under the Harvard Six Cities Study for a description of models and Table 2 for a list of covariates included in each model.

[c] Used 5-year age groups for stratification of baseline hazard function due to unsuitable risk estimates resulting from low numbers of deaths and large numbers of covariates.

Table 4. Relative Risks of Mortality from All Causes Associated with an Increase in Fine Particles for Selected Personal Characteristics in the Six Cities Study[a]

Characteristic	Percentage of Cohort	All-Cause Mortality
Age at Enrollment		
≤ 40	27.4	2.11 (0.88–5.07)
41–55	35.0	1.66 (1.17–2.35)
> 55	37.6	1.17 (0.98–1.40)
Gender		
Male	45	1.33 (1.08–1.63)
Female	55	1.20 (0.94–1.53)
Smoking Status		
Never-smoker	40	1.36 (1.02–1.82)
Former-smoker	24	1.29 (0.97–1.72)
Current-smoker	36	1.35 (1.04–1.74)
Education Level		
Less than high school	28	1.45 (1.13–1.85)
High school	38	1.30 (0.98–1.73)
More than high school	34	0.98 (0.72–1.36)
Occupational Exposure to Dust or Fumes[b]		
Yes	45	1.39 (1.13–1.72)
No	55	1.17 (0.92–1.50)
Marital Status		
Married	81	1.29 (1.08–1.54)
Other	19	1.42 (1.02–1.98)
Heart or Lung Disease[c]		
Yes	34	1.32 (1.06–1.63)
No	66	1.24 (0.99–1.57)
FEV_1[d]		
High	83	1.24 (1.03–1.49)
Low	17	1.35 (1.00–1.84)
FVC[d]		
High	85	1.28 (1.07–1.54)
Low	15	1.44 (1.02–2.02)
Community Influence[e]		
Not Harriman	85	1.28 (1.10–1.50)
Not Portage	80	1.26 (1.05–1.52)
Not Steubenville	83	1.31 (0.96–1.79)
Not St Louis	84	1.28 (1.10–1.50)
Not Topeka	85	1.28 (1.09–1.51)
Not Watertown	84	1.30 (1.11–1.53)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the Six Cities Study, this difference for fine particles was 18.6 $\mu g/m^3$. Analyses are based on the Extended Model with calendar year as the time axis and the baseline hazard function stratified by 1-year age groups and gender. See Table 2 for a complete list of covariates included in the Extended Model. Data are RRs with 95% CIs.

[b] Self-reported.

[c] Defined as doctor-diagnosed high blood pressure, heart disease, chronic bronchitis, emphysema, or asthma.

[d] High pulmonary function measure > 85% of predicted value based on subject's height and age. Low pulmonary function measure ≤ 85% of predicted value based on subject's height and age.

[e] Analysis dataset did not specify city.

Table 5. Relative Risks of Mortality from All Causes Associated with an Increase in Fine Particles for Selected Personal Characteristics and Education Level in the Six Cities Study[a]

	High School or Less		More Than High School	
Characteristic	*n*	All-Cause Mortality	*n*	All-Cause Mortality
Age at Enrollment				
≤ 40	1,189	2.42 (0.88–6.61)	1,035	0.87 (0.07–11.54)
41–55	1,895	1.70 (1.10–2.62)	942	1.30 (0.70–2.41)
> 55	2,273	1.32 (1.08–1.62)	777	0.86 (0.58–1.27)
Gender				
Male	2,330	1.48 (1.16–1.87)	1,341	1.07 (0.70–1.63)
Female	3,027	1.29 (0.97–1.70)	1,413	0.81 (0.49–1.36)
Smoking Status				
Never-smoker	2,099	1.65 (1.17–2.33)	1,174	0.88 (0.49–1.60)
Former-smoker	1,250	1.38 (0.99–1.94)	687	1.06 (0.54–2.09)
Current-smoker	2,008	1.38 (1.03–1.85)	893	1.02 (0.55–1.90)
Occupational Exposure to Dust or Fumes[b]				
Yes	2,722	1.49 (1.18–1.88)	923	1.11 (0.64–1.93)
No	2,635	1.31 (0.97–1.77)	1,831	0.88 (0.56–1.39)
Marital Status				
Married	4,336	1.42 (1.15–1.75)	2,237	0.96 (0.67–1.37)
Other	1,021	1.30 (0.89–1.90)	517	1.88 (0.63–5.61)
Heart or Lung Disease[c]				
Yes	1,940	1.48 (1.16–1.89)	828	0.95 (0.58–1.55)
No	3,417	1.28 (0.97–1.69)	1,926	1.17 (0.74–1.87)
FEV_1[d]				
High	4,361	1.34 (1.08–1.68)	2,398	0.95 (0.65–1.40)
Low	996	1.37 (0.97–1.94)	356	0.68 (0.25–1.86)
FVC[d]				
High	4,491	1.42 (1.15–1.76)	2,414	0.89 (0.61–1.31)
Low	866	1.45 (0.98–2.15)	340	1.21 (0.49–3.04)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the Six Cities Study, this difference for fine particles was 18.6 $\mu g/m^3$. Analyses are based on the Extended Model with calendar year as the time axis and the baseline hazard function stratified by 1-year age groups and gender. See Table 2 for a complete list of covariates included in the Extended Model. Data are RRs with 95% CIs.

[b] Self-reported.

[c] Defined as doctor-diagnosed high blood pressure, heart disease, chronic bronchitis, emphysema, or asthma.

[d] High pulmonary function measure > 85% of predicted value based on subject's height and age. Low pulmonary function measure ≤ 85% of predicted value based on subject's height and age.

the estimate of relative risk of the more educated group (95% CI: 0.94–15.35). None of the relative risks in the group with more than high school was statistically significantly different from unity ($P > 0.05$).

OCCUPATIONAL EXPOSURES

Occupational exposure to dusts, fumes, carcinogens, and other toxic substances is an important potential confounder in both of the studies under review because it is plausible that individuals who live in areas of high pollution tend, on average, to work in more polluted workplaces than subjects who live in clean areas. It is also plausible that subjects who work in polluted workplaces suffer higher risks of disease than subjects who work in clean workplaces. Indeed, there is extensive evidence that several workplace exposures (eg, asbestos, chromium) can cause lung cancer in workers (Siemiatycki et al 1991). (Credible estimates of the general population's attributable risk of lung cancer due to occupational exposures in industrialized countries are on the order of 10%.) There is also evidence that some workplace exposures can lead to nonmalignant respiratory disease (Christiani and Wegman 1995). For cardiovascular disease, however, although there are hints that a few workplace exposures may be risk factors, the evidence is weak and the attributable risk would be small. If there is an effect due to air pollution on any of these diseases, it is plausible that the effect differs depending on whether the subject has had significant occupational exposure to harmful substances in the workplace.

In both the Six Cities Study and the American Cancer Society (ACS) Study, some information was collected on the subjects' occupations and on their opinion as to whether they had been exposed to dusts and fumes in the workplace. This information had been used by the Original Investigators in their analyses to control for possible confounding by occupation. However, it is known that self-reported exposure to workplace substances is an inadequate indicator of exposure. Consequently, it is not clear that the self-reports of dusts and fumes and the simple white collar/blue collar variable created by the Original Investigators provided effective control for occupational confounders. The Reanalysis Team decided a more detailed assessment of the potential for confounding of the relation between particulate air pollution and mortality would be informative. The occupational data that were available in coded form were very limited. For the Six Cities Study, only the occupation and industry as recorded at the baseline interview were available.

Considering the type of data available and the nature of the diseases at issue, we developed a strategy to create two new variables that could be used to improve the control for possible confounding by workplace exposures. The first is a variable that we refer to as a "dirtiness index"; it describes, on a semiquantitative scale, the degree of dusts, gases, and fumes present in a subject's occupational environment. Conceptually, this is somewhat the same as assigning subjects to either white- or blue-collar worker categories. The dirtiness index plays a role similar to the "self-reported exposure to dusts, gases, and fumes" that had been used by the Original Investigators. We believe that the dirtiness index affords better control for general occupational exposures than either the self-reports by study subjects of exposure to dusts and fumes, or the Original Investigators' translation of job codes into a blue-collar/white-collar index. The lung carcinogen index was designed to indicate whether the subject's particular occupation would be considered to constitute an excess risk of lung cancer.

Occupational Exposure Indices

A research group within the Reanalysis Team that has had extensive and long-standing experience in assessing occupational exposure in the context of community-based studies (Gérin et al 1985; Siemiatycki et al 1991) oversaw the creation of new exposure indices. The development of these new indices of occupational exposure is described in detail in Appendix B (which is available from the Health Effects Institute upon request).

Briefly, the two new variables were based on the occupational/industrial coding systems that the Original Investigators had used, supplemented by additional information. In the case of the dirtiness index, the additional information came from work conducted in Montréal in the context of a large community-based cancer case-control study (Siemiatycki et al 1991). A dirtiness index had been developed and used in the Montréal study, and we adapted it to both the ACS Study and the Six Cities Study. For each of the 442 occupation codes in the 1970 US Census Classification system used to classify jobs in the Six Cities Study, we used the same criteria that had been used earlier in Montréal. With the resulting correspondences between job codes and dirtiness scores, the Reanalysis Team was able to attribute a measure of occupational dirtiness to each individual in the two studies. This index ranged from 0 (a very clean occupational environment) to 6 (a very dirty workplace environment).

In the case of the lung carcinogen indicator, the additional information came from lists of carcinogens evaluated by the International Agency for Research on Cancer (IARC), summarized by Boffetta and colleagues (1995), and by Ahrens and Merletti (1998).

Adjustment for Occupational Exposures

After calculating a dirtiness index score for job codes and assigning a binary variable for occupations with exposure to lung carcinogens (new occupational exposure indices; see Appendix B), we fit Cox proportional-hazards models identical to those that had been used by the Original Investigators, but with one or both of the new occupational covariates included in the models. We also carried out some analyses using the dirtiness index as a stratification variable to assess effect modification. We conducted all the analyses using calendar year as the time axis, as the Original Investigators had done, and we repeated them using age as the time axis. Because the resulting two sets of relative risks were virtually identical, we will present only the results using calendar time here.

Results

As shown in Appendix B, nearly 40% of all subjects were in the lowest (ie, cleanest) of the seven occupational dirtiness categories. The following population subgroups had much higher dirtiness levels than their respective complementary subgroups (see Table 6): males, subjects with less than high school education, and subjects who self-reported that they had exposure to dusts and fumes. Ever-smokers had slightly higher occupational dirtiness scores than never-smokers. Most importantly, subjects in Topeka and Watertown (among the least-polluted towns) had somewhat lower occupational dirtiness scores than subjects from other towns, and subjects in Steubenville were most likely to have jobs with high dirtiness scores. The percentage of subjects who worked in an occupation that has been shown or suspected to constitute an elevated risk of lung cancer was 7.5%. The patterns by gender, education, and smoking status for the indicator of occupational exposure to lung carcinogens were similar to those patterns observed for the dirtiness index. There was some variability by town of residence, but it was not clearly associated with the town's respective pollution level. There was some indication that cardiopulmonary disease and lung cancer were elevated in subjects who had higher dirtiness indices. Subjects who had ever been occupationally exposed to known lung carcinogens did not exhibit an elevated risk of lung cancer.

Table 7 shows estimates of the overall fine particle–mortality associations when different sets of covariates are included as confounders. In our reanalysis, neither the dirtiness index, in two different parameterizations, nor the lung carcinogen variable had any impact on the estimates of interest for all-cause mortality and cardiopulmonary disease mortality. For lung cancer mortality, the magnitude of the relative risk estimates was considerably reduced once the occupational confounders were included. Table 8 shows the relative risks of all-cause mortality, cardiopulmonary disease mortality, and lung cancer mortality due to air pollution among different subsets of the population. In contrast to the original analyses, in our statistical models we included the dirtiness index (as a categorical variable) for all three causes of death; in addition, for lung cancer mortality, we included the binary lung carcinogen variable. Cardiopulmonary mortality relative risks were distributed equally among males and females when all subjects were considered, and more heavily among never-smokers than ever-smokers. The lung cancer results were very unstable; there was an indication of greater effect of air pollution among males, especially among never-smokers, although deaths from lung cancer among the latter constituted a very small number of events.

Table 6. Occupational Dirtiness Scores and Prevalence of Occupational Exposure to Known Lung Carcinogens in the Harvard Six Cities Study

Characteristic	Mean Dirtiness Score[a]	Prevalence of Exposure to Lung Carcinogens (%)
All subjects	2.10	7.53
Air pollution by city		
Harriman	2.40	7.04
Portage	2.31	8.94
Steubenville	2.24	6.77
St Louis	2.31	9.27
Topeka	1.40	6.55
Watertown	1.85	6.13
Education level		
Less than high school	1.25	3.09
High school	2.10	8.46
More than high school	3.17	11.87
Occupational exposure to dust or fumes[b]		
Yes	2.85	10.17
No	1.46	5.31
Gender		
Female	1.72	5.49
Male	2.53	9.86
Smoker		
Never-smoker	1.90	7.96
Ever-smoker	2.23	6.87

[a] Occupational dirtiness score ranges from 0 (very clean) to 6 (very dirty) (SEs were all within the range 0.02–0.06).

[b] Self-reported.

Table 7. Relative Risks of Mortality from All Causes, Cardiopulmonary Disease, and Lung Cancer Associated with an Increase in Fine Particles Using the Original and Extended Models and Adjusting for Alternative Indices of Occupational Exposure in the Reanalysis of the Six Cities Study[a]

Model	All Causes	Cardiopulmonary Disease	Lung Cancer
Original[b]	1.26 (1.08–1.46)	1.31 (1.07–1.61)	1.40 (0.82–2.38)
Original + dirtiness A[c] (+ lung carcinogens[d])	1.24 (1.07–1.45)	1.28 (1.04–1.58)	1.32 (0.76–2.31)
Original + dirtiness B[e] (+ lung carcinogens)	1.27 (1.08–1.48)	1.34 (1.09–1.66)	1.30 (0.75–2.27)
Extended[f]	1.28 (1.09–1.49)	1.32 (1.07–1.63)	1.13 (0.65–1.97)
Extended + dirtiness A (+ lung carcinogens)	1.26 (1.07–1.47)	1.29 (1.04–1.60)	1.06 (0.59–1.91)
Extended + dirtiness B (+ lung carcinogens)	1.28 (1.09–1.50)	1.35 (1.09–1.68)	1.05 (0.59–1.89)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the Six Cities Study, this difference for fine particles was 18.6 $\mu g/m^3$. Data are RRs with 95% CIs.

[b] The Original Model included $PM_{2.5}$, indicators of current and former smoking, a two-level indicator of education, occupational exposure to dust or fumes, and body mass index; baseline hazard function was stratified by 1-year age groups. See Table 2 for a complete list of covariates included in the Original Model. For consistency with our Extended Model, occupational analyses using the Original Model are based on 1-year age stratification, rather than the 5-year age stratification used by the Original Investigators.

[c] Dirtiness A is a categorical dirtiness variable.

[d] A binary variable for occupations with exposure to lung carcinogens; used only in the analyses for lung cancer.

[e] Dirtiness B is a continuous dirtiness variable.

[f] The Extended Model included the following covariates: (1) the Original Model covariates except for current-smoker pack-years and the two-level indicator of education level; (2) current-smoker, years of smoking, cigarettes per day, indicators of age started smoking, a three-level indicator of education level, marital status, alcohol consumption; and (3) interactions between gender and each of three covariates: current-smoker, marital status, and alcohol consumption; baseline hazard function was stratified by 1-year age groups. See Table 2 for a complete list of covariates included in the Extended Model.

Table 8. Relative Risks of Mortality from All Causes, Cardiopulmonary Disease, and Lung Cancer Associated with an Increase in Fine Particles in Various Subsets of the Population Using the Original Model + Dirtiness + Lung Carcinogens in the Reanalysis of the Six Cities Study[a]

Group	All Causes	Cardiopulmonary Disease	Lung Cancer
All subjects	1.26 (1.08–1.48)	1.34 (1.08–1.65)	1.30 (0.75–2.27)
Females	1.19 (0.92–1.53)	1.33 (0.92–1.90)	0.67 (0.22–2.08)
Males	1.31 (1.07–1.61)	1.34 (1.03–1.75)	1.64 (0.85–3.16)
Never-smokers	1.24 (0.92–1.66)	1.39 (0.93–2.10)	3.88 (0.44–34.18)
Females	1.16 (0.80–1.67)	1.21 (0.70–2.08)	4.06 (0.46–36.12)
Males	1.25 (0.77–2.04)	1.61 (0.85–3.06)	NA[c]
Ever-smokers	1.33 (1.10–1.61)	1.37 (1.07–1.76)	1.40 (0.80–2.46)
Females	1.29 (0.90–1.84)	1.56 (0.96–2.54)	0.52 (0.13–2.10)
Males	1.38 (1.11–1.73)	1.33 (0.99–1.78)	1.82 (0.97–3.43)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the Six Cities Study, this difference for fine particles was 18.6 $\mu g/m^3$. The Original Model included the following covariates: $PM_{2.5}$, indicators of current- and former-smokers, current-smoker pack-years, former-smoker pack-years, a two-level indicator of education level, occupational exposure to dust or fumes, and body mass index. See Table 2 for a complete list of covariates included in the Original Model. "Dirtiness" is a continuous occupational variable; "lung carcinogens" is a binary variable for occupations with exposure to lung carcinogens and was used only in the analyses for lung cancer. Data are RRs with 95% CIs.

[b] The large upper confidence limit is due to the small number of deaths (8) in this group.

[c] NA = no deaths in that group.

Table 9. Relative Risks of Mortality from All Causes and Cardiopulmonary Disease Associated with an Increase in Fine Particles Using the Original Model Stratified by Occupational Dirtiness in the Reanalysis of the Six Cities Study[a]

Dirtiness	All Causes	Cardiopulmonary Disease
Low	1.28 (0.96–1.70)	1.49 (1.00–2.22)
Medium	1.08 (0.81–1.43)	1.19 (0.81–1.74)
High	1.47 (1.13–1.90)	1.45 (1.04–2.04)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the Six Cities Study, this difference for fine particles was 18.6 $\mu g/m^3$. See Table 2 for a complete list of covariates incorporated into the Original Model; baseline hazard function was stratified by 1-year age groups. For occupational dirtiness, low = 0, medium = 1–3, and high = 4–6. Data are RRs with 95% CIs.

Table 9 shows that the relative risks of mortality from air pollution differ by dirtiness stratum for all-cause mortality and cardiopulmonary disease mortality, but with no coherent trend; the lowest relative risk is in the middle dirtiness stratum. Table 10 shows the results of an analysis of the relative risks of air pollution for all-cause mortality stratified by dirtiness score and education level. There is no clear indication as to whether the air pollution effect is more dependent on occupational dirtiness or on education.

POPULATION MOBILITY

The Original Investigators in the Six Cities Study had examined the association between fine particle air pollution and mortality using a cross-sectional personal interview of subjects selected in six cities, with interviews conducted between 1974 and 1977. Although subjects had been reinterviewed 3, 6, and 12 years after the initial interview, and their residences were recorded during follow up, this information had not been used in the original analyses. Information on the number of years the subject had lived in the city of enrollment prior to recruitment also was recorded, but not used. Air pollution concentrations averaged over the follow-up period had been assigned to each individual by city regardless of the amount of time that individual had lived in the city of enrollment.

The Reanalysis Team attempted to evaluate the impact of population mobility, which would affect exposure to ambient air pollution, on mortality. Mobility both before and after enrollment in the study was considered.

Preenrollment Mobility

Only limited information was available on mobility within the cohort prior to enrollment. Partial residence histories, tied to job history, had been recorded on the initial questionnaire but not in computer files. However, the number of years in which subjects lived in the city of enrollment had been noted during the initial interview, and was available for analysis.

The distribution of the numbers of years subjects had resided in their community of enrollment before the study began is shown in Table 11 both by city and for all cities combined. Subjects had lived in the original city of enrollment for 30 years on average, ranging from an average of 23 years in Watertown to 44 years in St Louis. We note that the two of the most highly polluted cities (Steubenville and St Louis) also had the longest average residency of subjects prior to enrollment. When we included residency duration as a predictor of all-cause mortality, it did not change the association between fine particles and relative risk of mortality (RR = 1.28, 95% CI: 1.09–1.50); residency duration was a weak predictor of mortality (RR = 0.99 on the basis of the observed range of 74 years, 95% CI: 0.79–1.24). We obtained these results using the Extended Model with calendar year as the time axis.

We examined the potential for residency duration to modify the association between fine particles and mortality by relating fine particles to mortality within three levels of residency duration (< 20 years [34% of sample],

Table 10. Relative Risks of Mortality from All Causes Associated with an Increase in Fine Particles Using the Original Model Stratified by Occupational Dirtiness and Educational Level in the Reanalysis of the Six Cities Study[a]

Dirtiness	Less Than High School	High School	More Than High School
Low	1.72 (0.87–3.40)	1.40 (0.85–2.30)	0.94 (0.60–1.47)
Medium	0.97 (0.61–1.52)	1.13 (0.67–1.89)	1.26 (0.74–2.16)
High	1.67 (1.19–2.34)	1.65 (0.99–2.75)	0.93 (0.37–2.36)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the Six Cities Study, this difference for fine particles was 18.6 $\mu g/m^3$. See Table 2 for a complete list of covariates incorporated into the Original Model; baseline hazard function was stratified by 1-year age groups. For occupational dirtiness, low = 0, medium = 1–3, and high = 4–6. Data are RRs with 95% CIs.

Table 11. Distribution of Residence Duration Before Enrollment (in Years) by City of Enrollment in the Reanalysis of the Six Cities Study

			Percentiles						
City	Mean	SD	0	5	25	50	75	95	100
Harriman	24.9	16.2	0	3	13	22	33	58	74
Portage	25.9	18.1	0	3	9	24	38	60	74
Steubenville	36.0	16.4	0	8	24	37	48	62	73
St Louis	43.7	16.1	0	17	31	45	56	69	74
Topeka	25.7	15.6	2	5	13	24	35	55	74
Watertown	22.9	17.4	0	1	7	20	34	55	73
All cities	29.8	18.3	0	3	15	28	44	63	74

20–40 years [36%], and > 40 years [30%]). The relative risk of fine particles and all-cause mortality was 1.41 (95% CI: 0.94–2.12), 1.21 (95% CI: 0.91–1.62), and 1.32 (95% CI: 1.05–1.65), respectively, within these three groups. Consequently, the length of time spent in the community before enrollment does not appear to affect the association between fine particle air pollution and mortality.

Mobility After Enrollment

Subject mobility after enrollment had been ascertained through the use of annual letters, postcards, or phone calls to study participants. Follow-up interviews also had been conducted at 3, 6, and 12 years, which further extended the mobility database. The Reanalysis Team computerized this information for the purposes of assessing the influence of post-enrollment subject mobility on the association between air pollution and mortality.

A minority (18.5%) of the cohort had moved outside the city of enrollment before follow up was completed. Mobility increased with educational attainment; 12.8% of subjects with less than a high school education had moved, 16.9% of high school graduates had moved, and 25.0% of those subjects with more than high school education had moved. Mobility did not vary with occupational exposure to dust or fumes. Of those subjects not occupationally exposed, 19.2% had moved; 17.6% of those in the exposed group had moved. The frequency of moving was similar for all smoking status groups (19.7% for current-smokers, 16.8% for former-smokers, and 18.4% for never-smokers). Moving was less frequent among married persons (17.8%) than nonmarried persons (22.1%). Mobility was similar in males (18.2%) and females (18.7%). However, movers tended to be younger (average age at enrollment, 44.6 years) than nonmovers (50.8 years). Mobility was similar in all cities (12.7% to 19.0%) except Watertown (31.8%). The crude death rate (the number of deaths/number of subjects) was much lower for the movers (12.1%) compared with nonmovers (18.9%), likely due to the younger average age of subjects that moved.

Reanalysis showed that relative risk of fine particle exposure and all-cause mortality for the nonmover group was 1.30 (95% CI: 1.10–1.54), notably comparable to that for the entire sample (RR = 1.28, 95% CI: 1.09–1.49). We based this analysis on the Extended Model with calendar year as the time axis. The relative risk of movers was 1.08 (95% CI: 0.67–1.76), a value clearly lower than that observed for the nonmoving cohort. Subjects in the mover group tended to have higher educational attainment than did nonmovers. Fine particle pollution was not related to mortality in the group with higher education. We determined the relative risk within the three educational groups for movers and nonmovers separately. Among the nonmovers, the relative risk associated with fine particles was lower for the subjects with the highest level of education (RR = 1.41, 95% CI: 1.10–1.82 for subjects without high school education; RR = 1.42, 95% CI: 1.06–1.91 for subjects with high school education; and RR = 0.96, 95% CI: 0.68–1.35 for subjects with more than high school education). Our analysis showed a similar risk for subjects without high school education among the movers (RR = 1.56, 95% CI: 0.67–3.64) as for the nonmovers without high school education. However, we obtained relative risks less than 1 among the high school–educated movers (RR = 0.71, 95% CI: 0.26–1.99) and movers with more than high school education (RR = 0.96, 95% CI: 0.40–2.30). The weakness of the association between fine particles and mortality in the

mover group thus was due largely to those subjects with at least high school education.

The Reanalysis Team also conducted an analysis of population mobility in which subjects were treated as being lost to follow up once they moved out of the original city of residence. The advantage of this analysis is that subjects who moved are not assigned an inappropriate exposure level. The relative risk of fine particle exposure on mortality for this new analysis was 1.23 (95% CI: 1.05–1.45), a value only slightly lower than that observed for the entire cohort.

Finally, we conducted an analysis of the mover group using long-term average exposures to fine particles but ignoring follow-up data on this group before the time the subjects first moved from the city of enrollment. This analysis produced a relative risk of all-cause mortality of 1.25 (95% CI: 0.75–2.10), similar to that in the entire sample (RR = 1.28), but greater than that in our first analysis of the mover group (RR = 1.08) based on full follow-up information starting at the time of enrollment into the study. The confidence interval on estimates of the relative risk in the mover group is comparable to that in the entire sample. Our previous estimate of RR = 1.08 for the mover group based on full follow up may be biased low because some individuals who otherwise might have moved from the original city of residence may have died before they had the opportunity to do so. However, because members of the mover group were notably younger than members of the nonmover group, this bias is expected to be small.

TIME-DEPENDENT COVARIATES

The Reanalysis Team undertook Poisson regression analyses of data from the Six Cities Study to estimate the relative risk of mortality from fine particles while taking into account changes in the values of both the air pollution exposures and risk factors that occurred during follow up. The Cox proportional-hazards model used by the Original Investigators had provided an estimate of the relative risk under the assumption that exposure to fine particles remained fixed during follow up. Specifically, exposure to fine particles had been assigned by the Original Investigators using the mean exposure determined on the basis of samples taken between 1979 and 1985. In this section, we have used Poisson regression to provide a separate series of risk estimates that can be compared with those generated by the Original Investigators. More importantly, by using the Poisson model, we can evaluate the impact of temporal changes in the values of both fine particles and other risk factors.

Using fixed-in-time covariates, a positive association had been demonstrated between mortality and fine particle air pollution by the Original Investigators in the Six Cities Study, with the age-adjusted hazard ratio estimated from the Cox proportional-hazards model for the most-polluted city compared to that for the least-polluted being 1.26 (95% CI: 1.08–1.46). We also considered the potential confounding influence of several other variables measured at baseline: smoking status, number of pack-years of smoking, educational level achieved, and BMI.

During follow up of the Six Cities cohort, attempts had been made to reinterview subjects to ascertain changes in these covariates. Longitudinal data were available for up to four interview dates: date of enrollment, and 3, 6, and 12 years later. We evaluated the effects of changes in the values of these covariates over time using the Poisson regression model:

$$\log RR\,(z,w) = \log r\,(x,z,w) - \log r_0\,(x)$$

in which RR denotes the relative risk of mortality, z represents a set of covariates (BMI, education, smoking, and occupational exposure) that can modify the mortality rate r in addition to the effect of the air pollution exposure w, and x represents a set of covariates (here, age and gender) that describe the background mortality rate r_0. We fit this model to the Six Cities Study data using EPICURE (Preston et al 1993).

In order to compare results from the Poisson regression to previously derived relative risk estimates using the Cox proportional-hazards model, we first modeled exposure to fine particles by using a city-specific mean concentration of fine particles over the follow-up period. We assessed the effect of changes in exposure over time in later models that incorporated city-specific concentration levels calculated for the following periods: before 1981, 1981–1982, 1983–1984, 1985–1986, and 1987 or after. We calculated these values separately for each city by smoothing available mean annual levels of fine particles using log-linear regression.

We adjusted all models for gender and the following age groups: < 45, 45 through 49, 50 through 54, ... 75 through 79, and 80 or older; and also evaluated the effects of BMI, education, occupational exposures, and smoking. We categorized BMI into quartiles on the basis of frequency distribution in the study population at time of enrollment; specifically, we placed subjects into one of the following quartiles: < 22.70, 25.26, 28.21, and ≥ 28.21 kg/m^2. Using these same cutpoints, we also evaluated changes in BMI over time on the basis of data collected during follow-up interviews.

We created an indicator variable to denote whether or not an individual had completed high school education. Although a more detailed categorization of this variable was available, education was dichotomized to ensure

consistency with the approach that had been taken by the Original Investigators. Similarly, we assigned occupational exposure to dust and fumes using a binary variable.

The Reanalysis Team modeled the effect of smoking behavior on mortality three ways. First, we conducted an analysis by using the same variables as in the original study. These models included terms for current-smokers, former-smokers, cigarette pack-years for current-smokers, and cigarette pack-years for former-smokers. Thus we were able to examine whether Poisson regression (which we were using) produced results similar to Cox regression (which the Original Investigators had used). Second, we included in our model terms that represented the number of years of cigarette smoking (at baseline, or time of enrollment), and the number of packs of cigarettes smoked weekly. Finally, because of information obtained in follow-up interviews, we were able to model changes over time in the number of packs of cigarettes smoked weekly. (There were inconsistencies in the smoking status and number of smoking years reported during follow-up interviews, which precluded the use of these indicators of tobacco consumption as time-dependent covariates.)

The adjusted mortality rate ratios based on the Cox regression (Original Model) used by the Original Investigators (Table 12) provided a benchmark against which we compared similar estimates of risk generated by Poisson regression (Table 13). Both regression analyses are based on the same variables from the baseline questionnaire; however, unlike the Cox regression model, Poisson regression requires categorization of all variables, including BMI and cigarette consumption, prior to analysis. Nonetheless, there were no appreciable differences in the city-specific risk estimates obtained using the Cox and Poisson regression models. For example, the Poisson regression–based risk of mortality in Steubenville relative to that in Portage is 1.32 (95% CI: 1.11–1.57), comparable to the Cox regression–based relative risk of 1.26 (95% CI: 1.06–1.50).

Table 14 presents the relative risk estimates of mortality we obtained using the Cox and Poisson regression models with exposure to fine particles defined as a continuous covariate. Model 1 in Table 14 corresponds to the Original Model used by the Original Investigators. Model 2, which is based on Poisson regression with tobacco consumption as described in Model 1, gives slightly higher risk estimates than Model 1. Model 3, also based on Poisson regression but using duration and intensity of cigarette smoking at time of enrollment to characterize tobacco consumption, leads to risk estimates very close to those of Model 2. The agreement between Models 2 and 3 indicates that the two methods of controlling for tobacco consumption are equally effective.

Model 4 is the same as Model 3, except that the number of packs of cigarettes smoked per week is updated on the basis of information collected at the follow-up interviews 3, 6, and 12 years post-enrollment. Comparison of the relative risk estimates from these two models (RR = 1.31 and 1.32 for Models 3 and 4, respectively) indicates that the incorporation of time-dependent information on cigarette smoking did not have an appreciable impact on the association between particulate air pollution and mortality. Similarly, when we accounted for temporal changes in BMI (Model 5), it did not materially affect the relative risks for fine particles.

Table 15 shows the annual mean concentrations of fine particles between 1979 and 1988 within each of the six cities. Concentrations of fine particles decreased during the study period in Steubenville, Harriman, and St Louis; downward trends were less consistent in Portage, Topeka, and Watertown. The city-specific mean fine particle levels exhibited sizeable year-to-year variations.

Model 6 in Table 14 takes into account the generally declining levels of fine particles over time on the basis of the city-specific annual average fine particle concentrations shown in Table 15. The estimated relative risk of mortality associated with fine particles of 1.16 for Model 6 is lower than the comparable estimate of 1.31 for Model 5, although the confidence intervals for these two estimates demonstrate a degree of overlap.

There are several possible explanations for the attenuated relative risk estimates that were generated when fine particle exposures were modeled as calendar time-dependent variables (Model 6). First, it is possible that the smoothing of data using the log-linear regression did not yield exposures that were representative of those received by the residents in each city. Second, the use of time-dependent exposures resulted in less between-city variability in exposure to fine particles in the latter part of the follow-up period, during which most of the deaths had occurred; this lowered the relative risk of mortality per 18.6 $\mu g/m^3$ change in fine particle concentration. Finally, it is possible that for all-cause mortality, chronic exposure to fine particles is more important than acute exposure as a predictor of death. Unfortunately, we were unable to discriminate between risks of mortality estimated by using various exposure-time windows because of the high correlations between selected city-specific exposure indices based on various lag intervals.

AIR QUALITY DATA

A major strength of the Six Cities Study is that the Original Investigators had prospectively monitored a number of ambient air pollutants, using monitors specifically

Table 12. Adjusted Mortality Rate Ratios Estimated from Cox Proportional-Hazards Models: Original Results[a] from the Six Cities Study

Variable	All Subjects	Men	Women
Current-smoker	1.59 (1.31–1.92)	1.75 (1.32–2.32)	1.54 (1.16–2.04)
25 Pack-years of smoking	1.26 (1.16–1.38)	1.25 (1.12–1.39)	1.18 (1.00–1.41)
Former-smoker	1.20 (1.01–1.43)	1.17 (0.93–1.48)	1.34 (1.02–1.77)
20 Pack-years of smoking	1.15 (1.08–1.23)	1.16 (1.09–1.25)	1.15 (0.97–1.36)
Less than high school education	1.19 (1.06–1.33)	1.22 (1.06–1.41)	1.13 (0.95–1.35)
Body mass index	1.08 (1.02–1.14)	1.03 (0.95–1.12)	1.11 (1.03–1.20)
City[b]			
Portage	1.0	1.0	1.0
Topeka	1.01 (0.82–1.24)	1.04 (0.79–1.36)	0.97 (0.71–1.34)
Harriman	1.17 (0.97–1.41)	1.21 (0.96–1.54)	1.07 (0.79–1.45)
Watertown	1.07 (0.89–1.28)	0.94 (0.73–1.20)	1.22 (0.93–1.61)
St Louis	1.14 (0.96–1.36)	1.15 (0.91–1.44)	1.13 (0.86–1.50)
Steubenville	1.26 (1.06–1.50)	1.29 (1.03–1.62)	1.23 (0.93–1.61)

[a] Referred to as the Original Model by the Reanalysis Team; see Table 2 for a complete list of covariates incorporated into the Original Model. From Dockery et al 1993; corresponds to Table 2 in the original publication (Copyright © 1993, Massachusetts Medical Society, all rights reserved). Values are rate ratios (95% CIs). Rates have been adjusted for age, sex, and all other variables listed in the table. The rate ratios for body mass index are for an increase of 4.52 (1 SD). (Neither the text nor table in the original publication identify which pollutant is associated with these data.)

[b] City-specific rate ratios are all expressed in relation to Portage.

Table 13. Relative Risks of All-Cause Mortality in the Six Cities Study from Poisson Regression of Time-Varying Covariates[a]

Variable	All Subjects	Men	Women
Current-smoker	1.37 (0.98–1.87)	1.79 (1.05–2.86)	1.16 (0.73–1.74)
< 10 Pack-years[b]	1.0	1.0	1.0
10–30 Pack-years	1.57 (1.13–2.24)	1.35 (0.83–2.31)	1.74 (1.12–2.82)
> 30 Pack-years	1.87 (1.36–2.64)	1.56 (0.99–2.63)	1.93 (1.24–3.15)
Former-smoker	1.23 (0.99–1.52)	1.21 (0.88–1.63)	1.30 (0.95–1.75)
< 10 Pack-years[b]	1.0	1.0	1.0
10–25 Pack-years	0.96 (0.73–1.27)	0.96 (0.67–1.36)	1.12 (0.70–1.78)
> 25 Pack-years	1.47 (1.17–1.86)	1.54 (1.15–2.07)	1.73 (1.09–2.76)
Less than high school education	1.26 (1.13–1.41)	1.29 (1.12–1.49)	1.22 (1.03–1.45)
Body mass index[c]			
4th Quartile[b]	1.0	1.0	1.0
3rd Quartile	0.85 (0.74–0.98)	0.91 (0.77–1.09)	0.74 (0.59–0.93)
2nd Quartile	0.78 (0.67–0.90)	0.80 (0.66–0.97)	0.75 (0.59–0.93)
1st Quartile	0.82 (0.70–0.96)	0.98 (0.79–1.22)	0.69 (0.56–0.87)
City[d]			
Portage	1.0	1.0	1.0
Topeka	1.01 (0.82–1.24)	1.04 (0.79–1.36)	0.96 (0.69–1.31)
Harriman	1.16 (0.96–1.39)	1.20 (0.94–1.51)	1.06 (0.78–1.43)
Watertown	1.06 (0.89–1.27)	0.98 (0.77–1.24)	1.13 (0.86–1.49)
St Louis	1.13 (0.95–1.35)	1.16 (0.92–1.45)	1.07 (0.81–1.41)
Steubenville	1.32 (1.11–1.57)	1.39 (1.11–1.74)	1.22 (0.93–1.61)

[a] Risks have been adjusted for age, sex, and all other variables listed in this table.

[b] Other relative risks in this category are expressed in relation to this variable.

[c] Body mass index was categorized into quartiles based on the 8,111 subjects at baseline. Cutpoints in kg/m^2 were: ≤ 22.7, 25.26, 28.21, and > 28.21.

[d] City-specific relative risks are all expressed in relation to Portage.

Table 14. Relative Risks of Mortality from All Causes Associated with Selected Indices of Fine Particle Air Pollution[a] Based on Cox Proportional-Hazards Regression or Poisson Regression Models with Time-Dependent Covariates in the Reanalysis of the Six Cities Study

Model	Type	Covariates	Relative Risk (95% CI)
1	Cox	Age (5-year groupings), sex, current-smokers, pack-years for current-smokers, former-smokers, pack-years for former-smokers, high school education, body mass index, and occupational exposure to dust or fumes; values are based on data collected at baseline	1.26 (1.08–1.46)
2	Poisson	Same as Model 1[b]	1.32 (1.13–1.53)
3	Poisson	Age (5-year groupings), sex, number of years smoked, number of packs smoked per week, high school education, body mass index, and occupational exposure to dust or fumes; values are based on data collected at baseline	1.31 (1.13–1.53)
4	Poisson	Same as model 3 except deaths and person-years for category of "number of packs smoked per week" were calculated using changes indicated by follow-up interviews	1.32 (1.13–1.53)
5	Poisson	Same as model 4 except deaths and person-years for category of "body mass index" were calculated using changes indicated by follow-up interviews	1.31 (1.13–1.52)
6	Poisson	Same as model 5 except changes in exposure to particulate matter over time were incorporated into the model[c]	1.16 (1.02–1.32)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the Six Cities Study, this difference for fine particles was 18.6 $\mu g/m^3$. The exposure for each city was based on the mean of sampled measures taken between 1979 and 1985.

[b] The use of the Poisson regression model required the categorization of body mass index as well as duration, intensity, and cumulative tobacco consumption that had been modeled as continuous variables in the Cox model.

[c] Exposures were defined according to 13 calendar periods: earlier than 1979, 1979, 1980, 1981, ... , 1989, and 1990 or later.

Table 15. Annual Average Concentration of Fine Particles by Calendar Year in Each of the Six Cities[a]

Year	Harriman	Portage	Steubenville	St Louis	Topeka	Watertown
1979	—	11.4	40.3	24.0	12.6	16.7
1980	26.3	12.8	30.0	22.7	15.6	17.3
1981	20.7	11.4	33.5	19.9	15.1	16.3
1982	18.7	10.1	27.9	17.7	11.9	13.4
1983	19.5	11.4	25.4	17.3	11.8	12.3
1984	19.7	11.1	26.1	18.4	12.9	17.4
1985	20.1	9.3	24.7	18.0	10.5	14.5
1986	20.5	10.8	21.7	17.9	9.2	—
1987	18.6	10.7	28.6	—	10.7	—
1988	—	—	—	—	13.7	—
Mean (available years 1979–1985)	20.9	11.0	29.6	19.0	12.5	14.9
Mean (all available years)	20.7	10.9	28.7	18.7	12.1	14.9

[a] A dash (—) indicates that no fine particle data were collected for that year.

Table 16. Relative Risks of Mortality from All Causes, Cardiopulmonary Disease, and Lung Cancer Associated with Various Measures of Air Pollution from the Reanalysis of the Six Cities Study[a]

		Cause of Death		
Pollutant	Range[b]	All Causes	Cardiopulmonary Disease	Lung Cancer
$PM_{2.5}$	18.6 μg/m^3	1.28 (1.09–1.49)	1.32 (1.07–1.63)	1.17 (0.67–2.04)
SO_4^{2-}	8.0 μg/m^3	1.28 (1.09–1.50)	1.32 (1.06–1.63)	1.15 (0.66–2.01)
SO_4^{2-} adjusted[c]	9.1 μg/m^3	1.27 (1.09–1.48)	1.30 (1.05–1.59)	1.14 (0.66–1.96)
TSP	55.8 μg/m^3	1.26 (1.07–1.47)	1.21 (0.98–1.50)	1.25 (0.71–2.20)
PM_{15}	28.3 μg/m^3	1.28 (1.09–1.51)	1.30 (1.04–1.62)	1.21 (0.67–2.18)
H^+	25.8 nmol/m^3	1.12 (0.97–1.30)	1.25 (1.03–1.53)	0.97 (0.57,1.64)
SO_2	22.4 ppb	1.26 (1.08–1.47)	1.25 (1.01–1.54)	1.13 (0.66–1.95)
SO_2 reconstructed[d]	22.1 ppb	1.26 (1.08–1.48)	1.24 (1.00–1.54)	1.08 (0.63–1.88)
NO_2	15.8 ppb	1.25 (1.07–1.46)	1.28 (1.04–1.59)	1.15 (0.65–2.04)
O_3	8.3 ppb	0.87 (0.76–1.00)	0.78 (0.64–0.95)	0.94 (0.56–1.59)

[a] Data are RRs with 95% CIs.

[b] Unless otherwise noted, all ranges were calculated from the values in Table 17a in Part I of this report, which corresponds to Table 1 in Dockery et al 1993.

[c] This range was calculated by the Reanalysis Team to adjust for artifactual sulfate.

[d] This range was reconstructed by the Original Investigators during the reanalysis.

developed for this purpose. For the same study population used by the Original Investigators, the Reanalysis Team calculated relative risks for ambient air pollutants that had been measured in the Six Cities Study ($PM_{2.5}$, SO_4^{2-}, TSP, inhalable particles, H^+, SO_2, NO_2, and O_3). As indicated in Table 16, associations with all-cause mortality were demonstrated by a number of pollutants, including fine particles, sulfate particles, total suspended particles, inhalable particles, aerosol acidity, sulfur dioxide, and nitrogen dioxide. Of the pollutants they had measured, only ozone did not appear to be associated with all-cause mortality. With the exception of aerosol acidity, all pollutants that demonstrated an association with mortality yielded a relative risk comparable to that for fine particles (RR = 1.28, 95% CI: 1.09–1.49). However, as can be seen in Table 17, a high degree of multicollinearity is evident between the different pollutants measured in the Six Cities Study.

A higher relative risk of cardiopulmonary mortality (RR = 1.32, 95% CI: 1.07–1.63) than for all-cause mortality had been demonstrated by fine particles. As was the case with all-cause mortality, increased cardiopulmonary mortality was associated with all other pollutants except ozone. No significant association with lung cancer mortality was demonstrated by any of the pollutants measured in the Six Cities Study, although the relative risks for lung cancer mortality were greater than unity for all pollutants except aerosol acidity and ozone.

With only six cities and a single fixed-site monitor within each city, the Reanalysis Team did not attempt to fit multiple-pollutant models to these data to identify which of these pollutants were most strongly associated with mortality. Multiple-pollutant models were used, however, in the ACS Study, which included 151 cities in the sulfate cohort and 50 cities in the fine particle cohort (see the Spatial Analyses section).

During the course of the Part I audit, it became apparent that sulfate data collected between 1979 and 1984 had been obtained using high-volume samplers that were subject to a known artifact. As detailed in Part I, the Reanalysis Team constructed city-specific calibration equations to correct for this known artifact, and developed adjusted estimates of the city-specific sulfate levels in the Six Cities Study. (The original/corrected sulfate concentrations [μg/m^3] in the six cities were 8.1/7.9 in Harriman, 5.3/4.7 in Portage, 12.8/13.5 in Steubenville, 8.0/7.6 in St Louis, 4.8/4.4 in Topeka, and 6.5/5.9 in Watertown; see Table 14 in Part I.) The relative risk of mortality from all causes was, however, virtually unchanged (RR = 1.27, 95% CI: 1.09–1.48), when compared with the estimate calculated using the Extended

Table 17. Correlation Between Pollutants in the Six Cities Study

	$PM_{2.5}$	SO_4^{2-}	TSP	PM_{15}	H^+	SO_2	NO_2	O_3
$PM_{2.5}$	100	98	84	97	59	85	78	−53
SO_4^{2-}		100	83	94	50	85	78	−50
TSP			100	90	12	86	82	−36
PM_{15}				100	50	81	77	−43
H^+					100	17	32	−56
SO_2						100	84	−47
NO_2							100	−80
O_3								100

Model (RR = 1.28, 95% CI: 1.09–1.50; see Table 16), after adjustment for this artifact.

Although the Audit Team, during the Part I audit, was able to confirm the city-specific annual average air pollutant levels for most pollutants measured by the Original Investigators, the reconstructed results for sulfur dioxide were somewhat different from those originally reported. The largest difference occurred in the St Louis data, for which the reconstructed sulfur dioxide concentration of 9.2 ppb was notably lower than the original value of 14.1 ppb. Nevertheless, the Reanalysis Team, using the reconstructed sulfur dioxide concentrations, obtained a relative risk of all-cause mortality (RR = 1.26, 95% CI: 1.08–1.48) that was virtually identical to the relative risk calculated for the same study population used by the Original Investigators (RR = 1.26, 95% CI: 1.08–1.47; see Table 16).

FLEXIBLE MODELING

Two important assumptions lie behind the Six Cities Study's original analysis, which had been based on the Cox proportional-hazards model. First, the Cox proportional-hazards assumption requires that, for each variable in the model, the hazard ratio remains constant over the entire follow-up period. Second, as in all parametric general linear models, the effect of each continuous predictor on the log hazard is assumed to be linear. The Original Investigators had not reported on the validity of these assumptions in the context of the Six Cities Study data. The Reanalysis Team needed to verify these assumptions to ensure that the estimates of the effects of particulate air pollution, and other covariates, would be unbiased.

Evidence that these assumptions may not hold could offer new insights into the impact of particulate air pollution on mortality. The extent to which the hazard ratio for long-term exposure to particles remains constant over time is of particular interest in light of the changes in ambient fine particle concentration during the follow-up period. Verification of both assumptions for major potential confounders is important, because misspecification of the effects assumptions may result in residual confounding of the estimated association between exposure and mortality. For these reasons, we examined the proportional-hazards and linearity assumptions underlying the original analysis using a flexible spline regression model.

As described in Appendix C (available upon request from the Health Effects Institute), the regression spline modeling approach allows for the simultaneous flexible estimation of (1) changes over time in the log hazard ratios of interest, and (2) nonlinear effects of continuous independent variables. Simultaneous estimation and testing of both effects is essential because failure to account for nonlinearity may result in spurious evidence of time dependence, and vice versa. We modeled time-dependent effects using a quadratic spline with 5 degrees of freedom (*df*) and 4 *df* used to represent nonlinear effects.

To reduce the size of the dataset to tractable levels, we relied on separate analyses of four disjoint and complementary subsets of the entire cohort. Each subset included about 2,000 participants, selected by simple random sampling without replacement. To test the hypotheses of interest, we then combined the four subset-specific likelihood ratio test statistics and adjusted the degrees of freedom appropriately (see Appendix C for details). We stratified the analyses by sex and 5-year age groups, as in the original study, and adjusted the effect of particulates for current and former smoking and for BMI. We conducted sensitivity analyses by varying the degrees of freedom for the covariates, and by varying the set of covariates included in the model.

Our tests of the proportional-hazards assumption using the default 5 *df* regression spline model yielded marginally significant time-dependent effects for both fine particles ($P = 0.0320$) and sulfate ($P = 0.0316$). Sensitivity analyses indicated that the statistical significance of these effects was robust with respect to choice of the covariates in the model, and did not depend on whether the effect of particulate air pollution, at a given point in time, was constrained to be linear or not. In contrast, we found that the significance of the time-dependent effects depended strongly on the number of degrees of freedom used to model these effects. Whereas more flexible 4 *df* and 5 *df* models provided evidence of significant departures from the Cox proportional-hazards assumption, such departures were not significant with 3 *df* or less. (The latter, less flexible models fitted the data considerably less well.) This indicates that considerable flexibility is essential to detect time dependence of the adjusted effects of both types of particles.

Figures 2 and 3 show the 5 *df* quadratic spline estimates of the time-dependent log–hazard ratio for fine particles and sulfate, respectively. Both estimates suggest that the respective hazard ratio is a nonmonotone function of the follow-up time. Specifically, the impact of fine particles on the mortality hazard decreases to near zero after five years of follow up, but later increases to reach a peak at about 10 to 12 years of follow up. One possible explanation for this could be that the pattern of temporal changes in the fine particle effect may reflect concurrent changes in between-city variations of the yearly particle concentration levels. (Indeed, the middle graph in Figure 2 of the original publication by Dockery and colleagues [1993] shows a sharp increase in fine particle levels in Steubenville at about 11 years of follow up, which coincides with the peak in our Figure 2.)

Although yearly fine particle levels are not available for the first 5 years of the follow up in the Six Cities Study, it can be seen from the upper graph in Figure 1 of the original publication that TSP had decreased substantially during this period in the two cities with the highest air pollution levels. This suggests that fine particle levels also may have decreased during this period, which corresponds to the initial decrease in the Reanalysis Team's estimate of the time-dependent effect of fine particles (our Figure 2). Thus, both the initial decrease and later increase in the estimated impact of fine particles on mortality seem to coincide with concurrent changes in between-city differences in yearly fine particle levels. This suggests that estimation of the impact of air pollution on mortality may be refined by taking into account the yearly variation in particulate levels, as represented by time-dependent covariates. (In the Time-Dependent Covariates section, we present the results of an analysis of the relation between mortality and fine

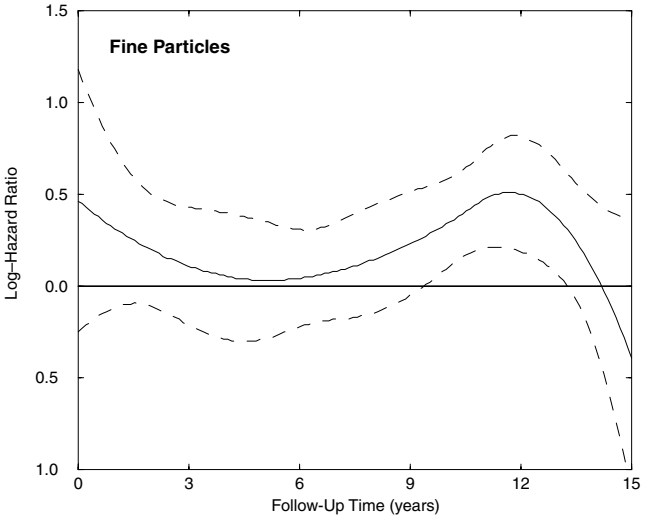

Figure 2. Change in the impact of fine particles over time in the Six Cities Study. Flexible quadratic spline estimate (5 *df*) of the time-dependent effect of fine particles on the log–hazard ratio of mortality in a subset of the Six Cities Study (n = 2,856 of which 1,430 were deaths). The log–hazard ratio was associated with a change in fine particles (18.6 μg/m^3) equal to the difference in mean concentrations between the most-polluted city and the least-polluted city. The solid curve represents the point estimate of the log–hazard ratio and the dashed curves represent the point-wise 95% confidence interval.

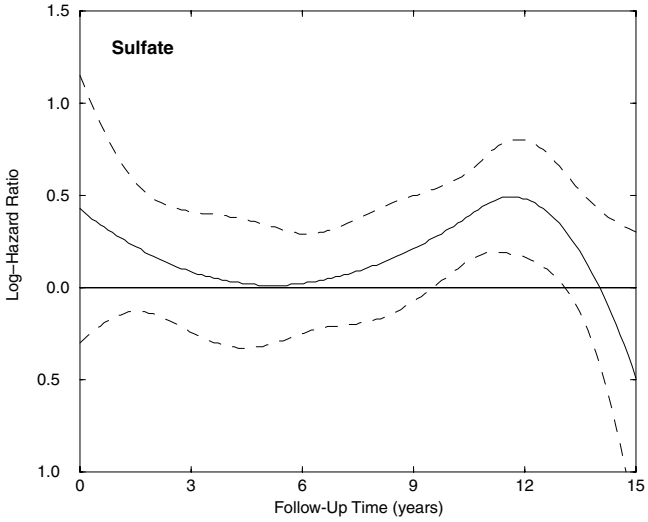

Figure 3. Change in the impact of sulfate over time in the Six Cities Study. Flexible quadratic spline estimate (5 *df*) of the time-dependent effect of sulfate on the log–hazard ratio of mortality in a subset of the Six Cities Study (n = 2,856 of which 1,430 were deaths). The log–hazard ratio was associated with a change in sulfate (8.0 μg/m^3) equal to the difference in mean concentrations between the most-polluted city and the least-polluted city. The solid curve represents the point estimate of the log–hazard ratio and the dashed curves represent the point-wise 95% confidence interval.

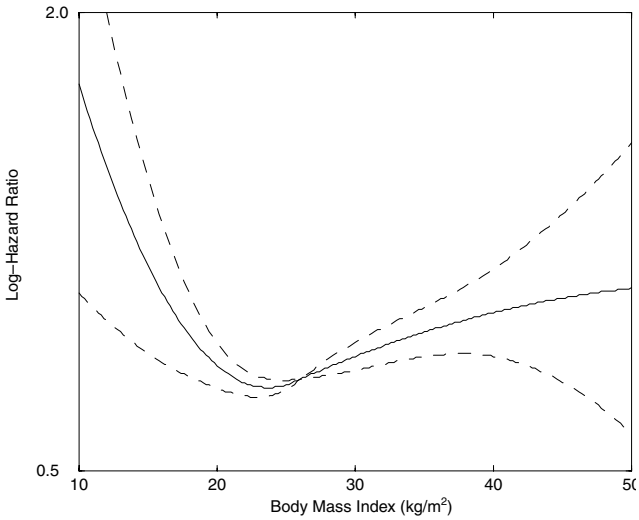

Figure 4. Flexible nonlinear estimate of the effect of BMI in the Six Cities Study. Flexible quadratic spline (4 *df*) estimate of the nonlinear effect of increasing BMI on the log–hazard ratio of mortality in a subset of the Six Cities Study (n = 2,856 of which 1,430 were deaths). The log–hazard ratio is plotted with respect to the mean BMI as the reference value. The solid curve represents the point estimate of the log–hazard ratio and the dashed curves represent the point-wise 95% confidence interval.

particle levels, using Poisson regression to take into account changes in fine particle levels over time.)

We found no evidence against the proportional-hazards assumption for BMI and the smoking variables considered ($P > 0.20$). However, we noted a significant departure from the linearity assumption for BMI. Figure 4 depicts the increases in the mortality hazard for both low and high values of BMI. This relation appears to be well approximated by the quadratic function used in the Full and Extended Models to characterize the effects of BMI.

AMERICAN CANCER SOCIETY STUDY

QUALITY ASSURANCE AUDIT OF THE DATA FOR THE AMERICAN CANCER SOCIETY STUDY

The Audit Team conducted a similar data audit of the ACS study, using data from the reduced ACS Cancer Prevention Study II (CPS-II) cohort described by Pope and colleagues (1995)*. There were three main differences between our audits of this and the Six Cities Study. First, the SAS data files that had been used in the original analysis were not available. Thus it was necessary for the ACS to reconstruct these datasets to correspond to the analytic files that had been used by the Original Investigators. Second, personnel who had been involved in the original formulation and conduct of CPS-II were no longer available to answer detailed questions about the procedures for data collection and management. Third, significant amounts of documentation for the ACS study were lost when ACS moved their main office from New York City to Atlanta. Thus, in comparison with the Six Cities Study, we had less documentation available to audit each variable; the auditable information and data were limited to microfilmed death certificates, microfilmed questionnaires, and some computer programming information. As was reported in Part I of this report, documentation of the ascertainment of vital status during the follow up no longer exists, nor does detailed information on the coding of each variable. Thus, the Audit Team often determined the coding rules by inference instead of documentation.

As we did for the Six Cities Study, we randomly selected 250 questionnaires and 250 death certificates for audit. We were able to trace microfilm copies of questionnaires and death certificates with the exception of three questionnaires (1.2%) and eight death certificates (3.2%). We were not able to decipher the causes of death on two additional death certificates.

Part II Audit

We audited variables for the Part II analysis by conducting a comparison of the data from baseline questionnaires to data in the electronic analysis file provided to the Reanalysis Team. Variables (in alphabetical order by SAS variable name from the analysis file) obtained from the baseline questionnaire and audited in Part II appear in Table 18.

We found no errors in 34 of the 55 audited variables. The error-free variables (by SAS name) were arthritis, asbestos, bladder disease, beer consumption (previous amount and years), chronic indigestion, cirrhosis of the liver, coal/stone dust and coal tar/pitch/asphalt exposure, colon polyps, breast cysts, diabetes, diverticulosis, diesel engine exhaust, duodenal ulcer, emphysema, exercise, formaldehyde exposure, gall stones, gynecologic problems, heart disease, heart medicine (two variables), prostate problems, rectal polyps, stroke, stomach ulcer, tuberculosis, thyroid medication, Tylenol (two variables), water additives, wine (previous years), and years resident in present neighborhood. Table 18 summarizes the errors we found in the remaining audited Part II variables.

Summary of Audit Findings

In this part of the audit, for the nonoccupational variables, we found no errors that would induce important effects (over 5%) in the statistical analyses; the highest error

* The original article appears in its entirety at the end of this Special Report.

Table 18. Findings from the Phase II Audit of the Original Study Questionnaires from the ACS Study

SAS Variable Name from the Analysis Files	Description of Variable	Number (and %) of Errors Found in 247 Questionnaires[a]	Comments from the Phase II Audit Team
ASTHMA	Asthma diagnosed by physician	1 (0.4)	Apparent coding error
COLDS	Colds/flu (number of times subject had colds or flu in the past year)	1 (0.4)	Apparent coding error
EVERSMK	Ever smoked cigarettes at least one per day for one year's time	5 (2.0	Apparent coding errors
HBP	High blood pressure diagnosed by physician	1 (0.4)	Apparent coding error
HEPTS	Hepatitis diagnosed by physician	1 (0.4)	Apparent coding error
HF	Hay fever diagnosed by physician	2 (0.8)	Apparent coding errors
KD	Kidney disease diagnosed by physician	2 (0.8)	Apparent coding errors
KS	Kidney stones diagnosed by physician	3 (1.2)	Apparent coding errors
LIQPR	Liquor (amount consumed in previous years)	2 (0.8)	Apparent coding errors
LIQPRYR	Liquor (years of previous consumption)	1 (0.4)	Apparent coding error
L_OCCUP	Last occupation/retired	18 (7.3)	Discussed in detail in Appendix A[b]
		103 (41.7)[c]	
MARITAL	Marital status	1 (0.4)	Apparent coding error
OCCUP	Occupation (current)	39 (15.8)	Discussed in detail in Appendix A
OCCUPYR	Occupation (total years in current occupation)	2 (0.8)	Apparent coding errors
OTH_JOB	Occupation (longest occupation)	20 (8.1)	Discussed in detail in Appendix A
OTH_YRS	Occupation (total years for longest occupation)	8 (3.2)	Discussed in detail in Appendix A
OTHER	Other medical conditions	2 (0.8)	Apparent coding errors
THYROID	Thyroid condition diagnosed by physician	1 (0.4)	Apparent coding error
THYRX	Thyroid medication (monthly consumption)	1 (0.4)	Apparent coding error
WATER	Water (source of drinking water)	3 (1.2)	Apparent coding errors
WINEPR	Wine (previous amount of consumption)	1 (0.4)	Apparent coding error

[a] Note that two questionnaires were missing and one copy of a questionnaire did not match the requested identification number.

[b] Appendix A is available on request from the Health Effects Institute.

[c] The analysis file contained entries for this variable that matched an adjacent, related column. If one interprets this variable without regard to the adjacent column, the error rate is 103/247 (41.7%); if one allows for this variation, the error rate is 18/247 (7.3%).

rate was 3.2%. However, we found very large discrepancies in the coding of occupation and industry, namely, last occupation/retired (error rate 7.3%), current occupation (15.8%), occupation of longest employment (8.1%), and total years of employment in longest occupation (3.2%). With the possible exception of the occupational data, our data quality audit results indicate that the information used in the ACS Study is of sufficient quality for use in the sensitivity analyses.

ALTERNATIVE RISK MODELS

The ACS Study Original Investigators' Analytic Approach

The association between ambient air quality and longevity had been examined by the Original Investigators in the ACS cohort using the Cox proportional-hazards model of survival. With this approach, the relative increase in the underlying hazard function, or instantaneous rate of death, was assumed to be modulated by a number of risk

factors for mortality, such as smoking habits, education, and air pollution, by a constant amount over the follow-up period. The time axis for this survival analysis was calendar year (1982 through 1989). Effects of age at enrollment into the study, gender, and race had been accounted for in the analysis by stratifying the baseline hazard function according to different categories of these covariates, with 5-year age groups used for age stratification. Other determinants of mortality that had been used by the Original Investigators in the Original Model are listed in Table 19.

In addition to overall mortality, cardiopulmonary disease, lung cancer, and all other causes excluding cardiopulmonary disease and lung cancer had been examined by the Original Investigators. Estimates of the log–relative risks had been obtained by maximizing the partial likelihood function for the Cox proportional-hazards model; and 95% confidence intervals for the log–relative risks had been obtained by adding and subtracting 1.96 times the standard error for the point estimate.

The Reanalysis Team's Analytic Approach

The Reanalysis Team considered a number of alternative risk models that included additional covariates not included in the Original Model, as well as different functional forms or categorizations of the original covariates. Also in our reanalyses, we used either calendar year or age as the time axis; when using age, we modeled age at enrollment into the study and age at event (death or censoring) in relation to air pollution and other determinants of mortality. This approach has been shown to more fully represent the effects of age on survival than does using calendar year as the time axis (Breslow and Day 1987).

We initially considered a Base Model that included air pollution with no additional determinants of mortality, with the baseline hazard function stratified by 1-year age groups, race, and gender. We then incorporated additional covariates into the Full Model used in the reanalysis (see Table 19). Specifically, we included square terms of continuous variables such as number of cigarettes smoked, years of smoking, and BMI in order to account for nonlinear effects on mortality. We also included variables to account for the age at which a subject started smoking and marital status, which had not been considered by the Original Investigators. To describe the effects of educational attainment in more detail, we considered three levels of education: less than high school, high school, and more than high school. We also included indicator variables for missing data on alcohol consumption due to the large fraction (nearly 70%) of missing observations; that is, 70% of the questionnaires did not have this information, which likely reflects a reluctance on the part of study participants to respond to this question. (A value of "no consumption" had been assigned to these missing data points by the Original Investigators.) We took into account the possibility that the effects of these risk factors could vary by gender by including an interaction term between gender and each of these factors.

We then developed a more parsimonious model by removing those variables that proved to be of least significance on the basis of Wald tests. We dropped a covariate from the Full Model if, when the covariate was removed, the P value based on the increase in the log-likelihood function was greater than 0.05. We continued this procedure until there was no statistical justification for removal of additional covariates. We did, however, keep a covariate in the model if the corresponding gender interaction was statistically significant ($P < 0.05$). We have referred to the parsimonious model derived in this way as the Extended Model (see Table 19).

The Reanalysis Team examined the potential effect of physical exercise on the relation between air pollution and mortality by including self-reported amounts of physical exercise (none or some, moderate, or heavy) as a covariate in the Extended Model. The level of physical exercise could be dependent on health status, with healthier people able to perform more intense exercise. Exposure to ambient air pollution also may increase the risk of developing disease; disease, in turn, may lead to less exercise. Thus, exercise level may be in the path of causation between exposure and death. We examined this possibility by including exercise level in the Extended Model for all causes of death for those people who reported ever having a selected number of diseases, and for those individuals who did not report having any of these diseases at the time of enrollment. These were defined as diseases or conditions for which a subject had ever been diagnosed by a doctor, and included high blood pressure, heart disease, stroke, diabetes, gall stones, chronic indigestion, kidney disease, kidney stones, bladder disease, cirrhosis of the liver, tuberculosis, chronic bronchitis, emphysema, asthma, stomach ulcer, duodenal ulcer, diverticulosis, rectal polyps, colon polyps, thyroid condition, arthritis, prostate trouble, and hepatitis.

Table 19. Covariates Included in the Original, Full, and Extended Models for the Reanalysis of the ACS Study[a]

Covariate	Alternative Risk Model		
	Original	Full	Extended
Tobacco consumption			
Current-smoker[b]	✓	✓	✓
Current-smoker years of smoking	✓	✓	✓
(Current-smoker years of smoking)^2		✓	✓
Current-smoker cigarettes per day	✓	✓	✓
(Current-smoker cigarettes per day)^2		✓	
Former-smoker years of smoking	✓	✓	✓
(Former-smoker years of smoking)^2		✓	✓
Former-smoker cigarettes per day	✓	✓	✓
(Former-smoker cigarettes per day)^2		✓	✓
Age started smoking (current-smoker) ≤18 years[b]		✓	✓
Age started smoking (current-smoker) > 18 years[b]		✓	✓
Age started smoking (former-smoker) ≤18 years[b]		✓	✓
Age started smoking (former-smoker) > 18 years[b]		✓	✓
Pipe and/or cigar smoker only[b]	✓	✓	✓
Passive cigarette exposure (hours/day)	✓	✓	✓
Education Level			
High school versus other[b]		✓	✓
More than high school versus other[b]		✓	✓
Less than high school versus other[b]	✓		
Occupational exposure[b,c]	✓	✓	✓
Body mass index	✓	✓	✓
(Body mass index)^2		✓	✓
Marital status			
Married versus single[b]		✓	✓
Other marital status versus married[b]		✓	✓
Alcohol consumption			
Drinks of alcohol per day	✓		
Beer consumption[b]		✓	✓
Missing beer consumption[b]		✓	✓
Wine consumption[b]		✓	✓
Missing wine consumption[b]		✓	✓
Liquor consumption[b]		✓	
Missing liquor consumption[b]		✓	

(Table continues next page)

[a] All three of these models were analyzed with standard Cox proportional-haards regressions.

[b] Dichotomous (yes/no) variable.

[c] Regular occupational exposure to any of the following: asbestos, chemicals/acids/solvents, coal or stone dusts, coal tar/pitch/asphalt, diesel engine exhaust, formaldehyde.

Table 19 *(continued)*. Covariates Included in the Original, Full, and Extended Models for the Reanalysis of the ACS Study[a]

Covariate	Alternative Risk Model		
	Original	Full	Extended
Interaction with gender			
Current-smoker[b]		✓	
Current-smoker years of smoking		✓	✓
(Current-smoker years of smoking)^2		✓	✓
Current-smoker cigarettes per day		✓	✓
(Current-smoker cigarettes per day)^2		✓	
Former-smoker years of smoking		✓	✓
(Former-smoker years of smoking)^2		✓	✓
Former-smoker cigarettes per day		✓	✓
(Former-smoker cigarettes per day)^2		✓	✓
Age started smoking (current-smoker) ≤ 18 years[b]		✓	
Age started smoking (current-smoker) > 18 years[b]		✓	
Age started smoking (former-smoker) ≤ 18 years[b]		✓	✓
Age started smoking (former-smoker) > 18 years[b]		✓	✓
Pipe and/or cigar smoker only[b]		✓	
Passive cigarette exposure (hours/day)		✓	
High school versus other[b]		✓	
More than high school versus other[b]		✓	
Less than high school versus other[b]		✓	
Occupational exposure to air toxics[b]		✓	✓
Body mass index		✓	✓
(Body mass index)^2		✓	✓
Married versus single[b]		✓	
Other marital status versus married[b]		✓	
Drinks of alcohol per day		✓	
Beer consumption[b]		✓	
Missing beer consumption[b]		✓	
Wine consumption[b]		✓	
Missing wine consumption[b]		✓	
Liquor consumption[b]		✓	
Missing liquor consumption[b]		✓	

[a] All three of these models were analyzed with standard Cox proportional-hazards regressions.

[b] Dichotomous (yes/no) variable.

[c] Regular occupational exposure to any of the following: asbestos, chemicals/acids/solvents, coal or stone dusts, coal tar/pitch/asphalt, diesel engine exhaust, formaldehyde.

Table 20. Relative Risks of Mortality by Cause of Death Associated with an Increase in Fine Particles or Sulfate in Risk Models with Alternative Time Axes in the Reanalysis of the ACS Study[a]

	Time Axis			
	Calendar Year		Age	
Alternative Risk Model[b]	Fine Particles	Sulfate	Fine Particles	Sulfate
All Causes [100%]				
Base	1.27 (1.18–1.37)	1.26 (1.19–1.33)	1.26 (1.17–1.35)	1.25 (1.18–1.32)
Original	1.18 (1.10–1.27)	1.16 (1.10–1.23)	1.18 (1.10–1.27)	1.16 (1.10–1.22)
Full	1.17 (1.09–1.26)	1.15 (1.08–1.21)	1.16 (1.08–1.25)	1.14 (1.07–1.20)
Extended	1.18 (1.09–1.26)	1.15 (1.09–1.21)	1.17 (1.09–1.25)	1.14 (1.08–1.20)
Cardiopulmonary Disease [50%]				
Base	1.41 (1.27–1.56)	1.39 (1.28–1.50)	1.41 (1.27–1.56)	1.38 (1.27–1.49)
Original	1.30 (1.18–1.45)	1.27 (1.17–1.38)	1.30 (1.18–1.45)	1.27 (1.17–1.37)
Full	1.28 (1.15–1.42)	1.25 (1.15–1.35)	1.28 (1.15–1.42)	1.24 (1.14–1.34)
Extended	1.30 (1.17–1.44)	1.25 (1.16–1.36)	1.29 (1.17–1.43)	1.25 (1.15–1.35)
Cardiovascular Disease [43%]				
Base	1.47 (1.32–1.65)	1.47 (1.35–1.60)	1.46 (1.31–1.63)	1.46 (1.34–1.59)
Original	1.36 (1.22–1.52)	1.36 (1.25–1.48)	1.36 (1.22–1.52)	1.35 (1.24–1.47)
Full	1.34 (1.20–1.49)	1.33 (1.22–1.45)	1.33 (1.19–1.48)	1.32 (1.21–1.43)
Extended	1.35 (1.21–1.51)	1.34 (1.23–1.46)	1.34 (1.20–1.50)	1.33 (1.22–1.44)
Respiratory Disease [7%]				
Base	1.07 (0.80–1.42)	0.94 (0.76–1.17)	1.09 (0.82–1.45)	0.95 (0.76–1.18)
Original	1.00 (0.76–1.33)	0.83 (0.67–1.04)	1.01 (0.76–1.34)	0.85 (0.68–1.05)
Full	0.96 (0.72–1.27)	0.81 (0.65–1.01)	0.99 (0.74–1.31)	0.82 (0.66–1.03)
Extended	0.98 (0.74–1.30)	0.82 (0.65–1.02)	1.00 (0.76–1.33)	0.83 (0.66–1.03)
Lung Cancer [8%]				
Base	1.23 (0.96–1.57)	1.63 (1.35–1.97)	1.21 (0.95–1.54)	1.62 (1.34–1.95)
Original	1.02 (0.80–1.29)	1.36 (1.13–1.65)	1.02 (0.80–1.30)	1.36 (1.12–1.64)
Full	0.99 (0.78–1.26)	1.32 (1.09–1.60)	0.98 (0.77–1.25)	1.31 (1.09–1.59)
Extended	1.00 (0.79–1.28)	1.33 (1.10–1.61)	0.99 (0.78–1.26)	1.32 (1.09–1.60)
Other Cancers [27%]				
Base	1.18 (1.03–1.36)	1.15 (1.03–1.28)	1.17 (1.02–1.34)	1.14 (1.02–1.26)
Original	1.14 (0.99–1.30)	1.10 (0.99–1.23)	1.13 (0.98–1.29)	1.10 (0.99–1.22)
Full	1.14 (1.00–1.31)	1.10 (0.99–1.23)	1.13 (0.98–1.29)	1.09 (0.98–1.21)
Extended	1.14 (0.99–1.31)	1.10 (0.99–1.22)	1.12 (0.98–1.29)	1.08 (0.97–1.21)
Other Causes [15%]				
Base	1.06 (0.88–1.27)	0.93 (0.81–1.08)	1.05 (0.88–1.26)	0.92 (0.80–1.06)
Original	1.01 (0.84–1.21)	0.88 (0.76–1.01)	1.01 (0.84–1.21)	0.87 (0.75–1.01)
Full	1.01 (0.84–1.21)	0.86 (0.75–1.00)	1.00 (0.83–1.20)	0.85 (0.74–0.99)
Extended	1.00 (0.84–1.21)	0.86 (0.75–1.00)	0.99 (0.83–1.19)	0.85 (0.74–0.99)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the ACS Study, this difference for fine particles was 24.5 $\mu g/m^3$, and for sulfate was 19.9 $\mu g/m^3$. Causes of death are shown with percentage of all causes. Data are RRs with 95% CIs.

[b] See the Alternative Risk Models section under the ACS Study for a description of models and Table 19 for a list of covariates included in each model.

Alternative Risk Estimates

The relative risks of mortality associated with increases in fine particles or sulfate evaluated at the ranges in exposure that had been considered by the Original Investigators are shown in Table 20 by covariate model specification (Base, Original, Full, and Extended), time axis used in the survival model (calendar year or age), and cause of death (all causes, cardiopulmonary diseases, cardiovascular diseases, respiratory diseases, lung cancer, other cancers, and all other causes). Compared with the Base Model with only air pollution, adjustment for selected risk factors for mortality reduced the relative risk associated with either fine particles or sulfate for all underlying causes of death for both time axes. We observed similar air pollution mortality risks in the three risk models with different groups of covariates: Original, Full, and Extended. The Full and Extended Models included terms for all gender interactions, age started smoking, and nonlinear (squared) terms for cigarettes smoked and BMI that had not been included in the Original Model. Although these additional covariates contributed to the overall characterization of the factors influencing mortality, their inclusion in the Full and Extended Models did not appreciably alter the asso-ciation between air pollution and mortality that had been observed in the Original Model.

Air pollution does not appear to be associated with either deaths from respiratory causes or the "other" causes of death, which include death from all causes except cardiopulmonary disease or cancer. However, both fine particles and sulfate are clearly associated with all-cause mortality and cardiovascular mortality. We found slightly higher air pollution risks if the underlying causes of death were restricted to ischemic heart disease (ICD-9 codes 410–414), with risks associated with sulfate of 1.32 (95% CI: 1.20–1.44) and risks associated with fine particle exposures of 1.37 (95% CI: 1.22–1.53), using the Extended Model with calendar year as the time axis. These results suggest that particulate air pollution may be affecting people with heart diseases more than it affects those with vascular problems.

Although the relative risk of death from lung cancer in relation to exposure to sulfate was significantly greater than unity, fine particles were not associated with an increased risk of lung cancer mortality. A weaker association was observed between deaths from other types of cancer and air pollution. Relative risks of mortality were similar in magnitude for fine particles and sulfate except in the case of lung cancer, for which the relative risk for fine particles (0.99) was less than that for sulfate (1.32). Reanalysis also showed that the association between air pollution and mortality was not sensitive to the specification of the time axis, suggesting that stratification of the hazard function by 1-year age groups was adequate to control for effects of age on survival.

Finally, to test the hypothesis that air pollution was not associated with a cause of death thought not to be affected by air pollution, we conducted an analysis of accidental mortality (ICD-9 codes > 800). Using the Extended Model with calendar year as the time axis, we estimated the risk of accidental deaths associated with particulate air pollution to be 1.07 (95% CI: 0.82–1.39) in the fine particle cohort and 1.01 (95% CI: 0.82–1.23) in the sulfate cohort, which confirmed our hypothesis.

Effect of Physical Activity and Disease History

We examined the effect of exercise on the association between ambient air pollution and mortality by including exercise level in the Extended Model for all causes of death. Of the full cohort, 28% reported no or a slight amount of exercise, 64% reported moderate exercise, and 8% reported heavy exercise. Exercise level was a determinant of mortality. For the sulfate cohort of 151 cities, for example, the relative risk of mortality associated with exposure to sulfate for subjects at the none/slight exercise level compared with those at the moderate exercise level was 0.63 (95% CI: 0.62–0.65); those engaged in heavy exercise had an even lower risk of mortality, with a relative risk of 0.54 (95% CI: 0.52–0.57). The inclusion of exercise in the Extended Model reduced the relative risk of sulfate from 1.15 (95% CI: 1.09–1.21) to 1.11 (95% CI: 1.05–1.18), using calendar year as the time axis, as it did for fine particles as well; the relative risk of mortality associated with fine particles was reduced from 1.18 to 1.13.

When we controlled for exercise level, the attenuation in risk associated with ambient air pollution was much less in the group without any reported diseases at time of enrollment (a reduction from 1.33 to 1.32 for sulfate and 1.30 to 1.29 for fine particles) than in the group with some reported disease (reduction from 1.14 to 1.10 for sulfate and 1.17 to 1.11 for fine particles). Although it was reduced somewhat, the air pollution effect persisted after we controlled for exercise. We found that the beneficial health effects of exercise were less obvious in the group without disease (RR = 0.88 for moderate versus none/slight, and 0.84 for heavy versus none/slight) than in the group with disease (0.63 for moderate and 0.53 for heavy exercisers).

We note that the effect of air pollution on mortality was more pronounced in people with no reported diseases than in the cohort with some reported disease. The group with no disease was younger overall than the group with a history of disease, with an average age at enrollment of

62.5 years compared to 66.9 years for the group with disease. Members of the disease-free group tended to die at an earlier age (66.8 years on average), compared with the group with disease (70.9 years), but experienced a lower percentage of deaths (3.4%) than the the group with disease (9.0%). Air pollution was associated with a relative risk of 1.30 in the disease-free group, with a corresponding increase of 1.0% in the number of deaths in this group. Air pollution also was associated with a relative risk of 1.14 for the group with disease, corresponding to a 1.3% increase in the number of deaths. The impact of air pollution on the number of deaths as a percentage of cohort members is thus seen to be greater in the group with a history of disease than in the disease-free group.

Shape of the Concentration-Response Function

The shape of the concentration-response function was examined by plotting city-specific estimates of the logarithm of the relative risk for each city compared with an index city against either fine particles or sulfate for three causes of death (all causes, cardiopulmonary disease, and lung cancer; Figure 5). We determined city-specific relative risk by including all individual risk factors in the Extended Model and the indicator functions for city (using one city as an index) from the Cox regression model, but excluding air pollution. For the sulfate cohort, we chose Greenville SC as the index city because it had sulfate levels near the overall mean concentration observed in the ACS Study. For the same reason, we selected Raleigh NC as the index city for the fine particle cohort of subjects. (Note, however, that we could have selected any index city with identical results.)

We didn't include data for Beaumont TX in these graphs for the sulfate cohort because the log–relative risk of this city was very low for all three causes of death. Boise City ID, with a fine particle concentration of 12.1 $\mu g/m^3$, also had a very low relative risk of all three causes of death compared with the index city. When we removed these two outlying data points, it enhanced the resolution of these graphs for assessing the shape of the concentration-response functions for fine particles and sulfate.

A nonparametric smoothed representation of the relation between air pollution and the city-specific logarithms of the adjusted relative risks is represented on each panel in Figure 6 using a locally weighted smoothing function (LOESS) with a 40% span (Cleveland and Devlin 1988), along with corresponding 95% confidence intervals.

The concentration-response function for sulfate demonstrates an increasing trend across the range of sulfate concentrations in the sulfate cohort, although the curve is relatively flat for concentrations of 10–15 $\mu g/m^3$. The concentration-response curves for fine particles and both all-cause and cardiopulmonary mortality demonstrate near-linear increasing trends through the range of particle levels observed in the fine particle cohort. The apparent absence of an association between lung cancer mortality and exposure to fine particles is consistent with our previous finding that the relative risk of lung cancer mortality was not elevated in this cohort. The relation between mortality and both fine particles and sulfate is explored further in the Flexible Modeling section using flexible spline regression models.

IDENTIFICATION OF SENSITIVE SUBGROUPS

In addition to examining the sensitivity of the association between air pollution and mortality to specifications of the risk model, the Reanalysis Team sought to identify population subgroups that may be especially susceptible to the health effects of exposure to air pollution. Unless otherwise specified, we have based all analyses of population subgroups on the Extended Model using calendar year as the time axis.

We examined effect modification by stratifying the cohort into categorical levels of the following covariates: educational attainment, reported heart or lung disease, exposure to air toxics, marital status, gender, smoking status (never-, former-, or current-smoker), exercise level, and age at enrollment. These stratified analyses, summarized in Table 21, permitted the Reanalysis Team to identify subgroups of the cohort that were more or less susceptible to sulfate or fine particle air pollution.

For both fine particles and sulfate, air pollution mortality risks decreased significantly ($P < 0.05$) with increasing educational attainment. We observed a similar pattern for cardiopulmonary disease and lung cancer causes of death (Figure 7). There was some evidence ($0.05 < P < 0.1$) that married persons demonstrated a reduced risk related to air pollution (RR = 1.14 for $PM_{2.5}$ and RR = 1.12 for SO_4^{2-}) compared with subjects who were not married at the time of the interview (RR = 1.31 for $PM_{2.5}$ and RR = 1.26 for SO_4^{2-}).

Although education appeared to be an effect modifier, it was not a strong confounder. The relative risk of mortality from all causes of death associated with exposure to sulfate was 1.16 (95% CI: 1.10–1.23), based on the Extended Model with no adjustment for education and calendar time as the time axis; we obtained a similar value after adjusting for education (RR = 1.15, 95% CI: 1.09–1.21). Exposure to fine particles yielded similar results (RR = 1.19 with no educational adjustment compared to 1.18 with adjustment for education) on the same basis. Education also was not a strong confounder of the air pollution effect for any of the

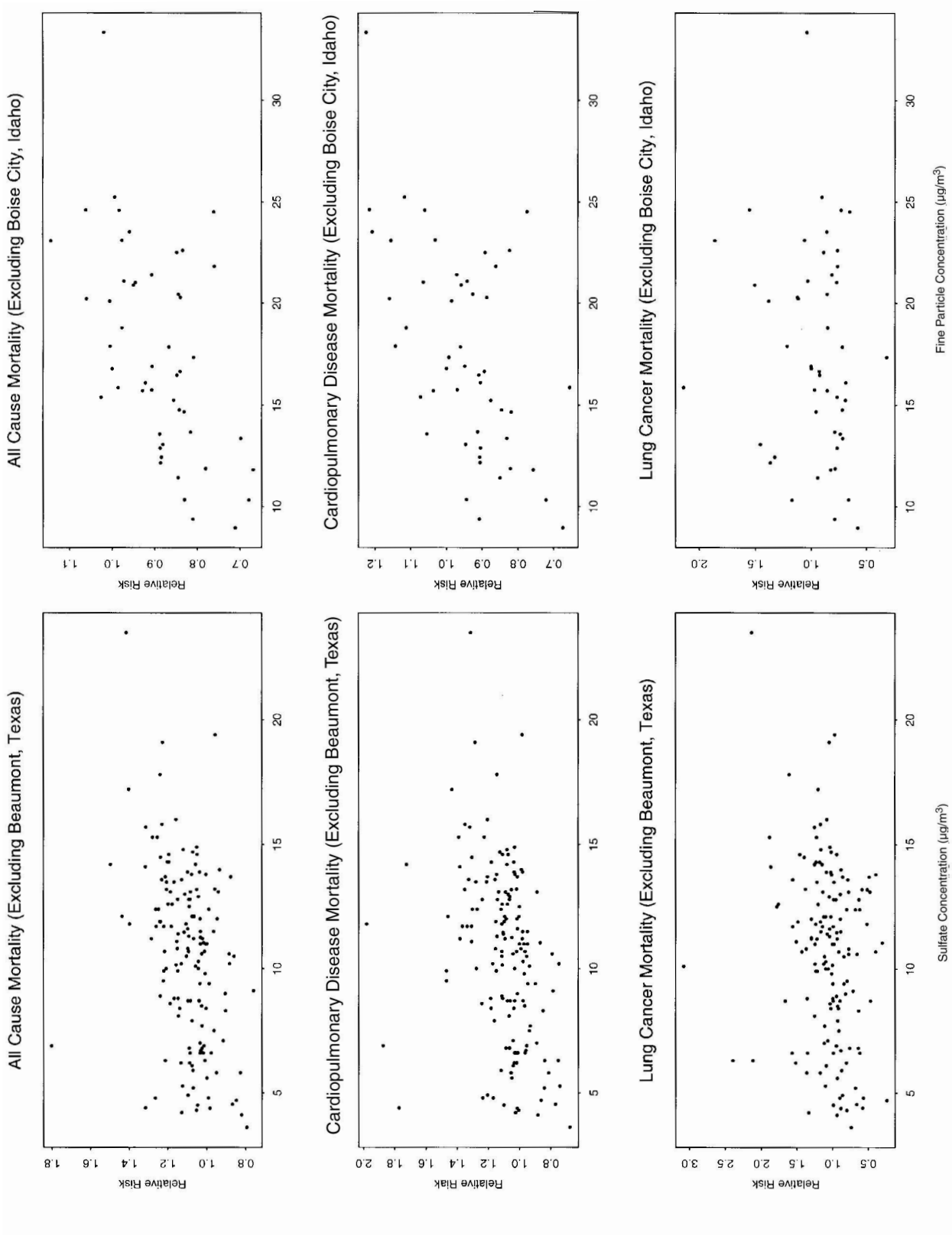

Figure 5. Shape of concentration-response function (relative risks) in the ACS Study. Relative risks of mortality from all causes, cardiopulmonary disease, and lung cancer by ambient concentrations of sulfate (linear-quadratic model) or fine particles (linear-quadratic-cubic model) for the reanalysis of the ACS Study. Based on the Extended Model and calendar year as the time axis. Relative risk scaled to unity at minimum concentration. Baseline hazard function stratified by 1-year age groups, gender, and race.

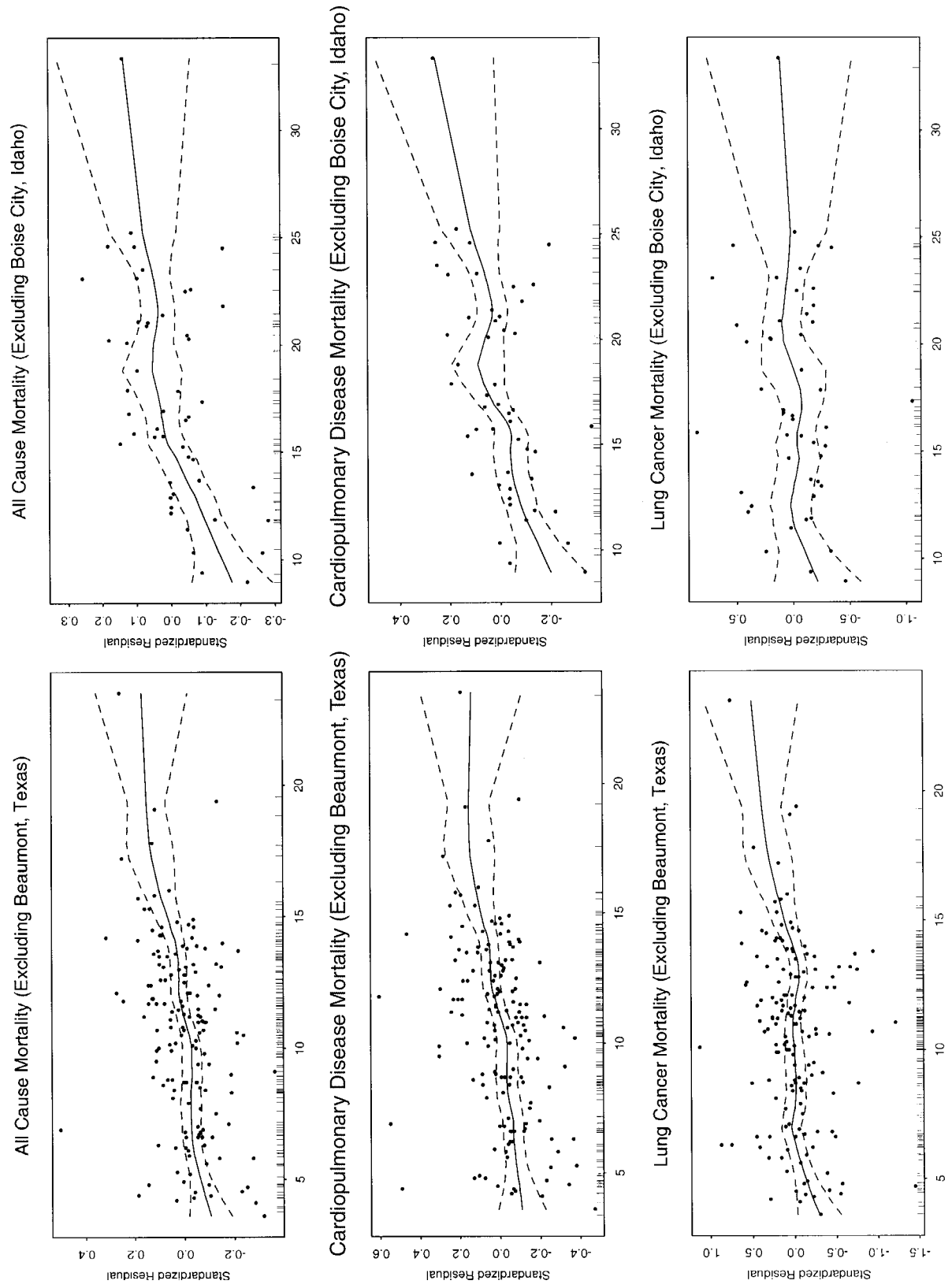

Figure 6. Shape of concentration-response function (standardized residuals) in the ACS Study. Standardized residuals of mortality from all causes, cardiopulmonary disease, and lung cancer by ambient concentrations of sulfate (linear-quadratic model) or fine particles (linear-quadratic-cubic model) for the reanalysis of the ACS Study. Based on the Extended Model and calendar year as the time axis. Standardized residual scaled to unity at minimum concentration. Baseline hazard function stratified by 1-year age groups, gender, and race.

Table 21. Relative Risks of Mortality from All Causes Associated with an Increase in Fine Particles or Sulfate for Selected Personal Characteristics in the ACS Study[a]

	Fine Particles		Sulfate	
Characteristic	Percent of Cohort	All-Cause Mortality	Percent of Cohort	All-Cause Mortality
Age at Enrollment				
< 50	29.3	1.19 (0.91–1.56)	29.3	1.14 (0.91–1.42)
50–60	36.4	1.13 (0.97–1.30)	36.5	1.12 (0.99–1.26)
> 60	34.3	1.19 (1.09–1.29)	34.2	1.16 (1.09–1.24)
Gender				
Male	43.6	1.17 (1.07–1.29)	43.4	1.12 (1.04–1.21)
Female	56.4	1.18 (1.06–1.32)	56.6	1.18 (1.09–1.29)
Smoking Status				
Never-smoker	48.4	1.25 (1.11–1.40)	48.3	1.18 (1.08–1.29)
Former-smoker	30.2	1.21 (1.07–1.37)	30.0	1.14 (1.03–1.25)
Current-smoker	21.4	1.14 (0.99–1.31)	21.7	1.21 (1.08–1.35)
Education Level				
Less than high school	11.3	1.35 (1.17–1.56)	12.3	1.27 (1.13–1.42)
High school	29.8	1.23 (1.07–1.40)	31.3	1.20 (1.08–1.33)
More than high school	58.9	1.06 (0.95–1.17)	56.3	1.05 (0.96–1.14)
Occupational Exposure to Dust or Fumes[b]				
Yes	19.4	1.08 (0.93–1.27)	19.8	1.14 (1.01–1.28)
No	80.6	1.20 (1.11–1.30)	80.2	1.15 (1.08–1.23)
Marital Status				
Married	84.0	1.14 (1.05–1.23)	84.0	1.12 (1.05–1.19)
Other	16.0	1.31 (1.13–1.52)	16.0	1.26 (1.12–1.41)
Disease Status[c]				
Heart or lung	37.1	1.15 (1.05–1.26)	37.2	1.15 (1.07–1.23)
Cancer	10.1	1.34 (1.15–1.57)	9.9	1.19 (1.05–1.34)
Other	63.7	1.19 (1.09–1.29)	63.2	1.12 (1.05–1.20)
Exercise				
No or slight	28.1	1.02 (0.90–1.15)	27.4	1.04 (0.95–1.15)
Moderate	64.4	1.19 (1.08–1.31)	64.7	1.16 (1.08–1.25)
Heavy	7.5	1.00 (0.73–1.37)	7.9	0.97 (0.76–1.23)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the ACS Study, this difference for fine particles was 24.5 $\mu g/m^3$, and for sulfate was 19.9 $\mu g/m^3$. Analyses are based on the Extended Model with calendar year as the time axis and the baseline hazard function stratified by 1-year age groups, gender, and race. See Table 19 for a complete list of covariates included in the Extended Model. Data are RRs with 95% CIs.

[b] Self-reported exposure to asbestos, chemicals/acids/solvents, coal or stone dust, coal tar/pitch/asphalt, diesel engine exhaust, or formaldehyde.

[c] Defined as doctor-diagnosed high blood pressure, heart disease, stroke, chronic bronchitis, emphysema, or asthma. Cancer defined as any type. Other diseases defined as diabetes, gall stones, chronic indigestion, kidney disease, kidney stones, bladder disease, cirrhosis of the liver, tuberculosis, stomach ulcer, duodenal ulcer, diverticulosis, rectal polyps, colon polyps, thyroid condition, arthritis, prostate trouble, or hepatitis.

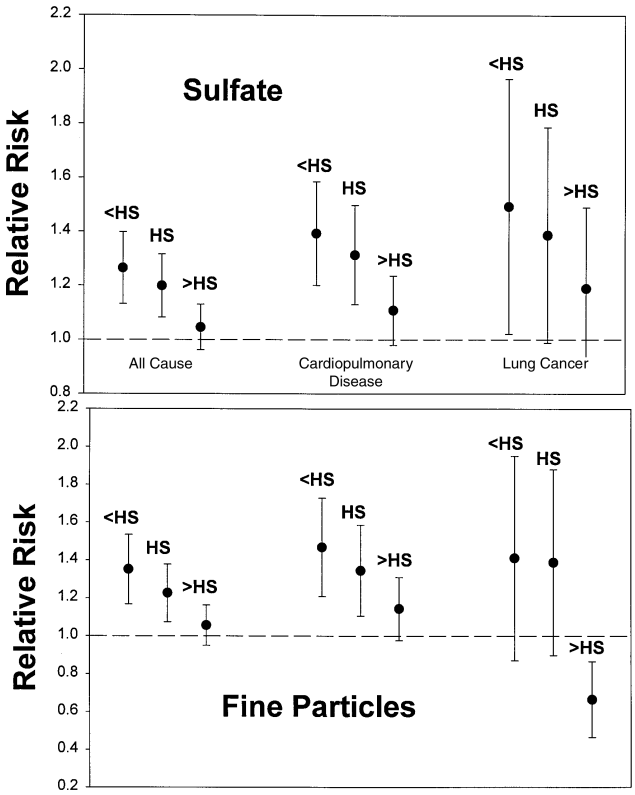

Figure 7. Relative risks of mortality by cause of death and educational attainment associated with sulfate or fine particles in the reanalysis of the ACS Study. HS = high school. Error bars represent ±2 SE.

specific underlying causes of death considered (results not shown).

The relative risk of lung cancer mortality for sulfate (RR = 1.33, 95% CI: 1.10–1.61) was greater than that for fine particles (RR = 1.00, 95% CI: 0.79–1.28). As shown in Figure 7, this difference in effect of air pollution largely can be explained by educational attainment. The relative risk of death from lung cancer associated with exposure to fine particles was 1.41 (95% CI: 0.87–2.29) for those individuals who had not completed high school, 1.39 (95% CI: 0.90–2.15) for those who had graduated from high school, and 0.66 (95% CI: 0.46–0.95) for those who had more than high school school education. The corresponding relative risks for exposure to sulfate were 1.49 (95% CI: 1.02–2.18), 1.39 (95% CI: 0.99–1.95), and 1.19 (95% CI: 0.89–1.59), respectively. The inverse relation between mortality and exposure to fine particles among individuals with more than high school education reduced the overall effect of fine particles on mortality; the relation between education and sulfate was attenuated by comparison. For subjects with high school education or less, the effects of fine particles and sulfate on lung cancer mortality were similar.

Although the general pattern of decreasing relative risk with increasing educational attainment shown in Figure 7 for all-cause and cardiopulmonary mortality is similar for fine particles and sulfate, the relative risk of lung cancer mortality is greater than unity (RR = 1.19) for sulfate and less than unity (RR = 0.66) for fine particles. In order to investigate the possibility that this difference might be due to the larger number of cities in the sulfate cohort (n = 151) than in the fine particle cohort (n = 50), we conducted a similar analysis restricted to the 47 cities for which both sulfate and fine particle measurements were available. This restricted analysis produced results similar to those obtained with the full sulfate and fine particle cohorts. Specifically, the relative risks of lung cancer mortality associated with sulfate based on the 47 cities were RR = 2.02 (95% CI: 1.25–3.25) for those with less than high school, 1.42 (95% CI: 0.91–2.21) for those with high school, and 1.14 (95% CI: 0.79–1.67) for those with more than high school education. The relative risks of lung cancer mortality associated with fine particles were 1.45 (95% CI: 0.89–2.36), 1.39 (95% CI: 0.89–2.16), and 0.72 (95% CI: 0.50–1.04), respectively, for the same three educational attainment groups.

To further characterize the effects of air pollution on mortality in relation to educational level or cohort, we classified members into two subgroups corresponding to high school education or less and more than high school education. For each of these two groups, we conducted analyses within categories of the following personal characteristics: exposure to air toxics, marital status, gender, smoking status, presence of heart or lung disease, exercise level, and age at enrollment. Table 22 illustrates that the relative risk for air pollution was greater in the lower education group than in the more educated cohort for all characteristics examined. On the basis of this analysis, it is not clear if there exists a subgroup of the cohort with more than high school education whose longevity is adversely affected by air pollution.

OCCUPATIONAL EXPOSURES

Occupational exposure is an important potential confounder in air pollution studies because it is plausible that individuals who live in highly polluted areas also work in more polluted environments. Extensive evidence indicates that several types of workplace exposure can cause lung cancer in exposed workers and can lead to nonmalignant respiratory disease. As described in the Occupational Exposures section of the Six Cities Study section, the Reanalysis Team attempted to control for occupational

Table 22. Relative Risks of Mortality from All Causes Associated with an Increase in Fine Particles or Sulfate for Selected Personal Characteristics by Educational Level and Sample Size[a] in the Reanalysis of the ACS Study[b]

	Fine Particles				Sulfate			
	High School or Less		More Than High School		High School or Less		More Than High School	
Characteristic	*n*	Relative Risk	*n*	Relative Risk	*n*	Relative Risk	*n*	Relative Risk
Age at Enrollment								
< 50	29,130	1.51 (1.00–2.27)	58,421	1.00 (0.70–1.43)	59,411	1.29 (0.92–1.81)	104,587	1.05 (0.78–1.41)
50–60	42,705	1.27 (1.02–1.60)	66,025	1.03 (0.84–1.26)	85,811	1.28 (1.08–1.52)	118,295	0.99 (0.83–1.17)
> 60	51,105	1.28 (1.14–1.43)	51,432	1.08 (0.95–1.23)	98,818	1.23 (1.13–1.34)	92,127	1.08 (0.97–1.19)
Gender								
Male	45,708	1.34 (1.17–1.52)	84,602	1.02 (0.89–1.17)	92,078	1.24 (1.12–1.37)	150,620	1.00 (0.90–1.12)
Female	77,231	1.24 (1.07–1.44)	91,276	1.12 (0.94–1.32)	151,962	1.24 (1.11–1.39)	164,389	1.12 (0.98–1.28)
Smoking Status								
Never-smoker	61,540	1.35 (1.15–1.57)	83,127	1.13 (0.95–1.35)	121,612	1.22 (1.09–1.37)	148,329	1.12 (0.98–1.28)
Former-smoker	29,191	1.40 (1.17–1.66)	34,680	1.06 (0.89–1.26)	63,596	1.25 (1.09–1.44)	104,016	1.03 (0.89–1.18)
Current-smoker	32,208	1.30 (1.08–1.57)	58,071	0.98 (0.80–1.21)	58,832	1.38 (1.19–1.59)	62,664	1.01 (0.85–1.20)
Occupational Exposure to Dust or Fumes[c]								
Yes	25,385	1.14 (0.93–1.39)	32,440	1.04 (0.81–1.32)	51,862	1.20 (1.03–1.39)	51,017	1.07 (0.88–1.30)
No	97,554	1.34 (1.20–1.50)	143,438	1.06 (0.94–1.19)	192,178	1.26 (1.15–1.37)	255,992	1.04 (0.95–1.14)
Marital Status								
Married	100,712	1.29 (1.15–1.44)	150,203	1.00 (0.89–1.13)	200,713	1.23 (1.12–1.34)	268,686	1.01 (0.92–1.11)
Other	22,227	1.32 (1.09–1.60)	25,675	1.29 (1.02–1.64)	43,327	1.30 (1.12–1.50)	46,323	1.19 (0.99–1.43)
Heart or Lung Disease[d]								
Yes	52,028	1.26 (1.11–1.42)	61,751	1.00 (0.87–1.15)	102,663	1.26 (1.15–1.38)	110,761	1.00 (0.90–1.29)
No	70,911	1.29 (1.09–1.53)	114,127	1.14 (0.97–1.35)	141,377	1.17 (1.03–1.32)	204,248	1.13 (0.99–1.29)
Exercise								
No or slight	30,840	1.08 (0.92–1.27)	51,984	0.94 (0.79–1.13)	59,538	1.12 (0.99–1.28)	91,533	0.94 (0.81–1.08)
Moderate	79,494	1.31 (1.15–1.49)	110,483	1.09 (0.95–1.25)	158,238	1.24 (1.12–1.37)	199,022	1.08 (0.97–1.20)
Heavy	10,642	1.40 (0.92–2.14)	11,711	0.69 (0.42–1.12)	22,317	1.05 (0.77–1.43)	21,417	0.80 (0.55–1.17)

[a] All *n* values include the two subcohorts of women who had been excluded from the original ACS analyses; however, they do not include subjects with missing data in a particular stratification variable.

[b] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the ACS Study, this difference for fine particles was 24.5 $\mu g/m^3$, and for sulfate was 19.9 $\mu g/m^3$. Analyses are based on the Extended Model with calendar year as the time axis and the baseline hazard function stratified by 1-year age groups, gender, and race. See Table 19 for a complete list of covariates included in the Extended Model. Data are RRs with 95% CIs.

[c] Self-reported exposure to asbestos, chemicals/acids/solvents, coal or stone dust, coal tar/pitch/asphalt, diesel engine exhaust, or formaldehyde.

[d] Defined as doctor-diagnosed high blood pressure, heart disease, stroke, chronic bronchitis, emphysema, or asthma.

confounding by supplementing the original datasets with two new variables: a dirtiness indicator and an indicator of exposure to occupational carcinogens.

The new exposure indices were created by a research team that has had extensive and long-standing experience in assessing occupational exposure in the context of community-based studies (Gérin et al 1985; Siemiatycki et al 1991). One index, the dirtiness indicator, was developed and used in Montréal in a large community-based cancer case-control study (Siemiatycki et al 1991). The other index, a lung carcinogen indicator, was developed with additional information provided by the International Agency for Research on Cancer (Boffetta et al 1995; Ahrens and Merletti 1998). During the baseline interview in the ACS Study, questions had been asked about current or last occupation and the occupation of longest duration. These had been coded by means of an ad hoc system developed by the ACS investigators. Whereas the Six Cities Study coding system had used 442 occupational and industrial categories, the ACS Study coding system had used only 68 occupational categories. Employing these codes, we allocated two new variables to each study subject.

Because the ACS Study used only occupation codes in a relatively small number of categories, it was often impossible

Table 23. Occupational Dirtiness Scores and Prevalence of Occupational Exposure to Known Lung Carcinogens in the ACS Study Sulfate Cohort

Characteristic	Mean Dirtiness Score[a]	Prevalence of Exposure to Known Lung Carcinogens (%)
All subjects	1.14	2.74
Air pollution[b]		
Low	1.17	2.92
Medium	1.12	2.6
High	1.13	2.74
Education level		
Less than high school	1.01	1.61
High school	1.13	3.68
More than high school	1.78	5.52
Occupational exposure to dust or fumes[c]		
No	0.91	1.55
Yes	0.08	7.57
Gender		
Female	0.69	0.25
Male	1.76	5.99
Smoker		
Never-smoker	1.03	1.87
Ever-smoker	1.24	3.55

[a] Occupational dirtiness score ranges from 0 (very clean) to 6 (very dirty) (SEs were less than 0.01).

[b] Based on tertiles of the distribution of sulfate across the 151 cities in the sulfate cohort.

[c] Self-reported in response to checklist of six occupational dusts and fumes.

to find a good fit between our occupation-industry combination and the ACS coding system. The occupation and industry codes used in the Six Cities Study allowed for a much better specification of at-risk jobs.

Appendix B shows that over half of all subjects were in the lowest (cleanest) of the seven occupational dirtiness categories. The following population subgroups had much higher dirtiness levels than their respective complementary subgroups (Table 23): males, subjects with less than high school education, and subjects who self-reported exposure to dusts and fumes. Smokers had slightly higher dirtiness scores than never-smokers or former-smokers. Most important, we found no clear relation between the occupational dirtiness scores and the pollution levels of the towns of residence.

The percentage of subjects who worked in an occupation that has been shown, or that is suspected, to be associated with an elevated risk of lung cancer was 2.7%. The patterns by subgroup were similar to those of the dirtiness index; again, we found no evidence of increasing exposure to occupational carcinogens with increasing environmental pollution.

As detailed in Appendix B, we found little evidence of any independent effect of the occupational dirtiness score on mortality from any of the causes examined. However, we found a relative risk of mortality from lung cancer associated with occupational exposure to lung carcinogens, as determined by our lung carcinogens variable, of 1.23 (95% CI: 1.00–1.51) in the fine particle cohort and 1.19 (95% CI: 1.02–1.39) in the sulfate cohort. Taken together, the lack of association between both new covariates and air pollution, and the equivocal findings on the associations between the new covariates and mortality, would suggest that the air pollution–mortality associations are unlikely to be confounded by either the occupational dirtiness score or the lung carcinogen variable.

Tables 24 and 25 show estimates of the overall air pollution–mortality associations when we included different sets of covariates as confounders. Neither the dirtiness index, in two different parameterizations, nor the lung carcinogen variable had a noticeable impact on the estimates of interest.

Table 26 shows the relative risks of mortality from all causes, cardiopulmonary disease, and lung cancer associated with sulfate and fine particle indices of air pollution among different population subgroups. In contrast with the original analysis, we included in the statistical models the dirtiness index for all three causes of death and, in addition, the lung carcinogen index for lung cancer mortality. We obtained results very similar to those that had been published by the Original Investigators.

We examined whether occupational dirtiness is an effect modifier for the air pollution effects. As indicated in Table 27, we found an apparently much greater effect of air pollution among subjects with the highest dirtiness score compared with those with low or medium levels of occupational dirtiness; however, there was no logical trend from the low to the medium category. We previously showed that educational attainment was also an important effect modifier, so we further explored the way the risk due to air pollution is mediated by education and occupational dirtiness. Table 28 shows the results of an analysis of the air pollution effect on all-cause mortality stratified by education level and dirtiness score. We see some indication in the fine particle cohort that the two effect modifiers have independent effects; this is less clear in the sulfate cohort, however, where the dirtiness variable appears to have a stronger impact than the education variable.

Table 24. Relative Risks of Mortality Associated with an Increase in Fine Particles Using the Original and Extended Models and Adjusting for Alternative Indices of Occupational Exposure in the Reanalysis of the ACS Study[a]

Model	All Causes	Cardiopulmonary Disease	Lung Cancer
Original[b]	1.19 (1.10–1.27)	1.30 (1.18–1.45)	1.03 (0.81–1.31)
Original + dirtiness A[c] (+ lung carcinogens[d])	1.18 (1.10–1.27)	1.30 (1.17–1.44)	1.03 (0.81–1.31)
Original[e]	1.16 (1.08–1.26)	1.29 (1.15–1.44)	1.02 (0.79–1.32)
Original + dirtiness B[f] (+ lung carcinogens)	1.16 (1.08–1.26)	1.29 (1.15–1.44)	1.02 (0.79–1.32)
Extended[g]	1.18 (1.09–1.26)	1.30 (1.17–1.44)	1.00 (0.79–1.28)
Extended[g] + dirtiness A (+ lung carcinogens)	1.17 (1.09–1.26)	1.29 (1.16–1.43)	1.00 (0.79–1.28)
Extended[e]	1.15 (1.07–1.24)	1.28 (1.15–1.43)	0.99 (0.77–1.28)
Extended[e] + dirtiness B (+ lung carcinogens)	1.15 (1.07–1.24)	1.28 (1.14–1.43)	0.99 (0.77–1.28)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the ACS Study, this difference for fine particles was 24.5 μg/m^3, and for sulfate was 19.9 μg/m^3. Data are RRs with 95% CIs.

[b] The Original Model included indicators of current and former smoking, a two-level indicator of education, occupational exposure to dust or fumes, body mass index, and an indicator of alcohol consumption; baseline hazard function was stratified by 5-year age groups. See Table 19 for a complete list of covariates included in the Original Model. This analysis used a cohort of 298,817 subjects. For consistancy with our Extended Model, occupational analyses using the Original Model are based on 1-year age stratificaton, rather than the 5-year age stratification used by the Original Investigators.

[c] Dirtiness A is a categorical dirtiness variable.

[d] A binary variable for occupations with exposure to lung carcinogens; used only in the analyses for lung cancer.

[e] This analysis used only the 274,022 observations for which occupational codes were available.

[f] Dirtiness B is a continuous dirtiness variable.

[g] The Extended Model included the Original Model covariates plus other indicators of smoking status, a different two-level indicator of education, marital status, other indicators of alcohol consumption, and interactions between gender and various other covariates; baseline hazard function was stratified by 1-year age groups. See Table 19 for a complete list of covariates included in the Extended Model. This analysis used a cohort of 298,817 subjects.

Table 25. Relative Risks of Mortality Associated with an Increase in Sulfate Using the Original and Extended Models and Adjusting for Alternative Indices of Occupational Exposure in the Reanalysis of the ACS Study[a]

Model	All Causes	Cardiopulmonary Disease	Lung Cancer
Original[b]	1.17 (1.10–1.23)	1.27 (1.17–1.38)	1.36 (1.13–1.65)
Original + dirtiness A[c] (+ lung carcinogens[d])	1.16 (1.10–1.23)	1.26 (1.17–1.40)	1.36 (1.12–1.64)
Original[e]	1.14 (1.08–1.21)	1.25 (1.15–1.36)	1.34 (1.09–1.64)
Original + dirtiness B[f] (+ lung carcinogens)	1.14 (1.07–1.21)	1.25 (1.15–1.36)	1.34 (1.09–1.64)
Extended[g]	1.15 (1.09–1.21)	1.25 (1.16–1.36)	1.33 (1.10–1.61)
Extended[g] + dirtiness A (+ lung carcinogens)	1.14 (1.08–1.21)	1.25 (1.15–1.35)	1.32 (1.09–1.60)
Extended[e]	1.12 (1.06–1.19)	1.24 (1.13–1.35)	1.31 (1.07–1.60)
Extended[e] + dirtiness B (+ lung carcinogens)	1.12 (1.06–1.19)	1.23 (1.13–1.34)	1.31 (1.07–1.61)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the ACS Study, this difference for fine particles was 24.5 μg/m^3, and for sulfate was 19.9 μg/m^3. Data are RRs with 95% CIs.

[b] The Original Model included indicators of current and former smoking, a two-level indicator of education, occupational exposure to dust or fumes, body mass index, and an indicator of alcohol consumption; baseline hazard function was stratified by 1-year age groups. See Table 19 for a complete list of covariates included in the Original Model. This analysis used a cohort of 559,049 subjects.

[c] Dirtiness A is a categorical dirtiness variable.

[d] A binary variable for occupations with exposure to lung carcinogens; used only in the analyses for lung cancer.

[e] This analysis used only the 511,031 observations for which occupational codes were available.

[f] Dirtiness B is a continuous dirtiness variable.

[g] The Extended Model included the Original Model covariates plus other indicators of smoking status, a different two-level indicator of education, marital status, other indicators of alcohol consumption, and interactions between gender and various other covariates; baseline hazard function was stratified by 1-year age groups. See Table 19 for a complete list of covariates included in the Extended Model. This analysis used a cohort of 559,049 subjects.

Table 26. Relative Risks of Mortality from All Causes, Cardiopulmonary Disease, and Lung Cancer Associated with an Increase in Fine Particles or Sulfate in Various Subsets of the Population Using the Original Model + Dirtiness + Lung Carcinogens in the Reanalysis of the ACS Study[a]

Model	All Causes	Cardiopulmonary Disease	Lung Cancer
Fine Particles			
All subjects	1.16 (1.08–1.26)	1.29 (1.15–1.44)	1.02 (0.79–1.32)
Females	1.14 (1.02–1.29)	1.40 (1.17–1.67)	0.83 (0.53–1.28)
Males	1.18 (1.06–1.30)	1.22 (1.06–1.40)	1.14 (0.83–1.56)
Never-smokers	1.23 (1.08–1.40)	1.36 (1.13–1.64)	0.75 (0.29–1.90)
Females	1.21 (1.03–1.42)	1.44 (1.14–1.83)	0.74 (0.25–2.21)
Males	1.29 (1.02–1.62)	1.21 (0.88–1.65)	0.85 (0.14–5.10)
Ever-smokers	1.13 (1.03–1.24)	1.24 (1.08–1.42)	1.03 (0.79–1.35)
Females	1.07 (0.90–1.28)	1.33 (1.01–1.76)	0.83 (0.51–1.33)
Males	1.15 (1.03–1.29)	1.21 (1.04–1.42)	1.14 (0.83–1.58)
Sulfate			
All subjects	1.14 (1.07–1.21)	1.25 (1.15–1.36)	1.34 (1.09–1.64)
Females	1.18 (1.07–1.29)	1.36 (1.18–1.56)	1.23 (0.86–1.75)
Males	1.11 (1.03–1.20)	1.19 (1.06–1.32)	1.39 (1.09–1.79)
Never-smokers	1.19 (1.08–1.31)	1.33 (1.15–1.53)	2.08 (1.03–4.23)
Females	1.20 (1.06–1.35)	1.34 (1.12–1.60)	2.15 (0.92–5.03)
Males	1.17 (0.99–1.39)	1.29 (1.03–1.62)	2.03 (0.56–7.33)
Ever-smokers	1.12 (1.04–1.21)	1.21 (1.09–1.35)	1.28 (1.04–1.58)
Females	1.15 (1.00–1.32)	1.38 (1.11–1.72)	1.08 (0.74–1.59)
Males	1.11 (1.01–1.21)	1.16 (1.03–1.31)	1.38 (1.07–1.78)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the ACS Study, this difference for fine particles was 24.5 $\mu g/m^3$, and for sulfate was 19.9 $\mu g/m^3$. The Original Model included the following covariates: $PM_{2.5}$, indicators of current- and former-smokers, current-smoker pack-years, former-smoker pack-years, a two-level indicator of education level, occupational exposure to dust or fumes, and body mass index; baseline hazard function was stratified by 1-year age groups. See Table 19 for a complete list of covariates included in the Original Model. "Dirtiness" is a continuous occupational variable; "lung carcinogens" is a binary variable for occupations with exposure to lung carcinogens and was used only in the analyses for lung cancer. Data are RRs with 95% CIs.

Table 27. Relative Risks of Mortality from All Causes and Cardiopulmonary Disease Associated with an Increase in Fine Particles or Sulfate Using the Original Model Stratified by Occupational Dirtiness in the Reanalysis of the ACS Study[a]

	Fine Particles		Sulfate	
Dirtiness	All Causes	Cardiopulmonary Disease	All Causes	Cardiopulmonary Disease
Low	1.13 (1.02–1.25)	1.28 (1.10–1.48)	1.12 (1.03–1.21)	1.25 (1.11–1.40)
Medium	1.10 (0.96–1.27)	1.13 (0.92–1.40)	1.10 (0.98–1.22)	1.12 (0.95–1.31)
High	1.39 (1.16–1.67)	1.52 (1.18–1.96)	1.30 (1.14–1.49)	1.46 (1.21–1.76)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the ACS Study, this difference for fine particles was 24.5 $\mu g/m^3$, and for sulfate was 19.9 $\mu g/m^3$. Analyses based on the Original Model; baseline hazard function was stratified by 1-year age groups. See Table 19 for a complete list of covariates incorporated into the Original Model. For occupational dirtiness, low = 0, medium = 1–3, and high = 4–6. Data are RRs with 95% CIs.

Table 28. Relative Risks of Mortality from All Causes Associated with an Increase in Fine Particles or Sulfate Using the Original Model[a] Stratified by Occupational Dirtiness and Educational Level in the Reanalysis of the ACS Study[b]

Dirtiness	Less Than High School	High School	More Than High School
Fine Particles			
Low	1.26 (1.01–1.57)	1.13 (0.94–1.36)	1.08 (0.93–1.25)
Medium	1.48 (1.03–2.13)	1.12 (0.81–1.54)	1.02 (0.85–1.23)
High	1.46 (1.09–1.95)	1.42 (1.03–1.95)	1.41 (0.98–2.03)
Sulfate			
Low	1.41 (1.18–1.67)	1.09 (0.95–1.26)	1.01 (0.89–1.14)
Medium	1.11 (0.85–1.46)	1.16 (0.91–1.48)	1.07 (0.93–1.23)
High	1.20 (0.97–1.48)	1.50 (1.19–1.90)	1.32 (1.00–1.74)

[a] Analyses based on the Original Model without the two-level indicator of education; baseline hazard function was stratified by 1-year age groups. See Table 19 for a complete list of covariates incorporated into the Original Model.

[b] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the ACS Study, this difference for fine particles was 24.5 $\mu g/m^3$, and for sulfate was 19.9 $\mu g/m^3$. For occupational dirtiness, low = 0, medium = 1–3, and high = 4–6. Data are RRs with 95% CIs.

ALTERNATIVE AIR QUALITY DATA

The Original Investigators' Approach

A database had been developed by the Original Investigators on sulfate particle concentrations [SO_4^{2-}(OI)] from high-volume samplers for total suspended particles, starting with sulfate data for 98 cities assembled by Özkaynak and Thurston (1987), who had previously developed city-specific ambient sulfate levels using monitoring data from the National Aerometric Database (NAD). Specifically, annual average sulfate concentrations had been calculated by Özkaynak and Thurston for 1980 using monitoring stations that met selection criteria established by the US EPA. This database then was augmented by the addition of average annual sulfate concentrations for 27 cities not meeting EPA's criteria for annual coverage, bringing the total number of cities to 127.

In order to further increase the number of cities available for analyses, sulfate data obtained from the Inhalable Particle Monitoring Network (IPMN) had also been included by the Original Investigators. Sulfate data from the IPMN were used for an additional 29 cities. Sites were required to have reported at least ten samples in each quarter of the year, and the annual mean had been computed as the mean of the four quarterly means. Sulfate means thus were available for a total of 154 metropolitan statistical areas (MSAs). Of these, three MSAs were not represented in the ACS cohort, which left 151 cities for analysis by the Original Investigators.

Data on fine particles from the IPMN had been obtained by the Original Investigators for 50 cities as reported in Lipfert and colleagues (1988). Because only median values by city were displayed in this report, median rather than mean values were used to characterize annual fine particle concentrations in the analysis conducted by the Original Investigators.

The Reanalysis Team's Approach

In order to evaluate the sensitivity of the original findings to the indicators of exposure to fine particle air pollution used by the Original Investigators, the Reanalysis Team constructed a number of alternative indicators of ambient particle levels using data from the US EPA's Aerometric Information Retrieval System (AIRS). We also obtained 1980 to 1989 daily 24-hour cumulative concentrations of TSP, and sulfate from TSP, for all monitoring stations in as many of the cities used in the original ACS analysis as possible. This latter information was extracted from the AIRS database by the Center for Air Pollution Impact and Trend Analysis (CAPITA) at Washington University in St Louis.

Sulfate data derived from TSP were available from AIRS for 132 cities in 1980, 124 in 1981, and no more than 60 cities in 1982 to 1989. Because of the marked reduction in sulfate monitoring in the latter period, we restricted our attention to the two years 1980 and 1981 for which data were available for at least 124 cities. In addition to ambient sulfate concentrations, we also retrieved supplementary data on land use surrounding the monitor (mobile, commercial, residential, agricultural, or industrial). In further sensitivity analyses, we restricted our attention to sulfate data for which there were at least 20 observations per year among all monitoring stations within a given city. Imposition of this selection criterion resulted in 107 eligible cities

for 1980, and 111 cities for 1981. There were a total of 126 cities for which sulfate concentrations were available for either 1980 or 1981. There were a total of 156 cities for which TSP data were available for either 1980 or 1981.

The high-volume samplers employed in the National Aerometric Database used glass-fiber filters, which were subject to artifacts because sulfur dioxide was present in the atmosphere. The sulfate measurements obtained from the IPMN were not subject to such artifacts because Teflon filters were used. In 41 cities both monitoring systems were employed. We adjusted the sulfate values obtained using glass-fiber filters to those obtained using Teflon filters by applying a linear regression equation in which city-specific averages were compared. We present details of the methods used and equations formed in Appendix D (which is available on request from the Health Effects Institute). We also developed separate calibration equations for three regions of the United States (West; Ohio Valley and Northeast; and East) and two seasons (April to September and October to March), because both sulfate and sulfur dioxide levels vary by region and season. We then augmented the city-specific average adjusted sulfate values, for those cities without sulfate observations from AIRS, by average sulfate values from the IPMN. This resulted in estimates for 144 of the 151 cities that had been examined by the Original Investigators. We were unable to find sulfate data for seven cities in either the AIRS or IPMN databases.

Recognizing that artifactual sulfate is associated with the use of glass-fiber filters in air quality samplers, the Reanalysis Team conducted an analysis of the association between mortality and ambient sulfate after correcting for the artifactual sulfate using a calibration equation we developed empirically. To compare our results with those obtained by the Original Investigators, however, we conducted other sensitivity analyses using the uncorrected sulfate data. (The results presented below permit one to assess the impact of the artifactual sulfate on the sulfate-mortality associations.)

The Reanalysis Team also obtained data from the IPMN directly from the EPA (this network is maintained by EPA for research rather than monitoring purposes). For the data pertinent to the ACS Study, the network consisted of dichotomous (DC) samplers with 15-μm and 2.5-μm cutpoints that measured PM_{15}(DC) (the mean inhalable fraction from dichotomous samplers), $PM_{15-2.5}$(DC) (mean coarse fraction from dichotomous samplers), and $PM_{2.5}$(DC) (mean fine fraction from dichotomous samplers). The IPMN also maintained high-volume samplers measuring mass, or total suspended particles [TSP(IPMN)], in addition to high-volume samplers using size-selective inlet (SSI) technology to record PM_{15}(SSI). Each method and instrument that measured mass also recorded data on sulfate concentrations.

Table 29 presents the definitions of pollutant variables and the sources of pollutant data. The city-specific mean or median levels of each of the indices of fine particle air pollution developed by the Reanalysis Team are presented in Appendix D. These values formed the basis for the following sensitivity analyses.

Risk Estimates Based on Alternative Air Quality Data

The means or medians of various indices of air pollution are summarized in Table 30. The median fine particle concentrations that had been used by the Original Investigators are denoted by $PM_{2.5}$(OI MD). These values are in good agreement with $PM_{2.5}$(DC MD), the median fine particle concentrations based on data from the dichotomous samplers, used by the Reanalysis Team, and are slightly less than the mean values $PM_{2.5}$(DC) used by the Team. Note that the sulfate levels SO_4^{2-}(OI) that had been used by the Original Investigators on the basis of 1980 monitoring data are comparable to the unadjusted sulfate data for the years 1980–1981 inclusive [$SO_4^{2-}{}_{(cb\text{-}unadj)}$] used by the Reanalysis Team. Adjustment by region and season for the artifactual sulfate resulted in notably reduced mean sulfate levels for $SO_4^{2-}{}_{(cb\text{-}adj\ US)}$, $SO_4^{2-}{}_{(cb\text{-}adj\ region)}$, and $SO_4^{2-}{}_{(cb\text{-}adj\ season)}$.

Figure 8 shows a comparison of the city-specific median concentrations of fine particles used by the Original Investigators and by the Reanalysis Team. With the exception of results for Denver CO, these two datasets had very good agreement. We calculated a median value of 7.20 μg/m^3 for fine particles in Denver, whereas the Original Investigators had used a median value of 16.09 μg/m^3. The IPMN database used by the Reanalysis Team included two stations operating in Denver from July 1980 to June 1983, which yielded median values of 5.67 μg/m^3 and 15.39 μg/m^3, respectively. A third station, which operated from July 1980 to March 1983, recorded a median value of 8.75 μg/m^3. A fourth station operated as a duplicate colocated station from July 1981 to June 1982, yielding a median value of 9.31 μg/m^3. In the absence of more detailed information on the source of the values reported by Lipfert and colleagues (1988), it is not possible to resolve this discrepancy between the values that had been used by the Original Investigators and those calculated by the Reanalysis Team.

We evaluated the influence of this discrepancy on the association between mortality and fine particle air pollution by removing the data for Denver from the analysis. We determined that Denver had not been an influential

Table 29. Summary of Pollutant Variables and the Sources of Data Used in the Reanalysis of the ACS Study

Pollutant	Definition	Source of Data	Number of Cities in Original ACS Dataset[a]	Number of Cities in Alternative Dataset	Used by
$PM_{2.5}$(DC)	**Mean** fine particle fraction	Dichotomous samplers; based on IPMN 1979–1983		63	Reanalysis Team
$PM_{2.5}$(DC MD)	**Median** fine particle mass concentration	Dichotomous samplers; based on IPMN 1979–1983		50	Reanalysis Team
$PM_{2.5}$(OI MD)	**Median** fine particle mass concentration	Based on IPMN 1979–1983	50		Original Investigators
$PM_{15-2.5}$(DC)	**Mean** coarse particle fraction (15-μm particles minus 2.5-μm particles)	Dichotomous samplers; based on IPMN 1979–1983		63	Reanalysis Team
PM_{15}(DC)	**Mean** inhalable particle fraction	Dichotomous samplers; based on IPMN 1979–1983		63	Reanalysis Team
PM_{15}(SSI)	**Mean** inhalable particle fraction	High-volume SSI samplers; based on IPMN 1979–1983.		59	Reanalysis Team
SO_4^{2-}(OI)	Sulfate data	Based on NAD 1980–1981	151		Original Investigators
SO_4^{2-}(DC)	Sulfate data from PM_{15}(DC)	Based on IPMN 1979–1983		51	Reanalysis Team
$SO_4^{2-}{}_{(cb\text{-}unadj)}$	Sulfate data for 1980–1981 inclusive, unadjusted for artifactual sulfate	Based on IPMN and NAD 1980–1981		144	Reanalysis Team
$SO_4^{2-}{}_{(cb\text{-}adj\ season)}$	Sulfate data for 1980–1981 inclusive, with season-specific adjustment for artifactual sulfate	Based on IPMN and NAD 1980–1981		144	Reanalysis Team
$SO_4^{2-}{}_{(cb\text{-}adj\ region)}$	Sulfate data for 1980–1981 inclusive, with region-specific adjustment for artifactual sulfate	Based on IPMN and NAD 1980–1981		144	Reanalysis Team
$SO_4^{2-}{}_{(cb\text{-}adj\ US)}$	Sulfate data for 1980–1981 inclusive, with US-specific adjustment for artifactual sulfate	Based on IPMN and NAD 1980–1981		144	Reanalysis Team
TSP	Total suspended particles	High-volume samplers; based on NAD 1980–1981		156	Reanalysis Team
TSP(IPMN)	Mean total suspended particle mass concentrations	High-volume samplers measuring mass TSP; based on IPMN 1979–1983		58	Reanalysis Team

[a] Of the 50 cities for which fine particle data were available, only 3 did not also have sulfate data available; therefore, a total of 154 cities contributed air quality data for the ACS Study.

observation in the dataset used by the Original Investigators, as neither the distribution of the fine particle data (Table 30) nor the relative risks of mortality (Table 31) varied appreciably with the inclusion or exclusion of the Denver data. However, when we used our fine particle data (eg, 1.14 with Denver and 1.17 without Denver for all-cause mortality) the relative risks for mortality were slightly reduced.

For IPMN data, $PM_{2.5}$(DC) was correlated weakly with $PM_{15-2.5}$(DC) ($r = 0.11$), but associated more strongly with PM_{15}(DC) ($r = 0.65$). Sulfate, however, was associated positively with fine particles ($r = 0.53$). Sulfate values that had been developed by the Original Investigators were correlated highly with those developed by the Reanalysis Team ($r = 0.92$) for the 144 cities with values in common (Figure 9). The distributions of the two measures of sulfate, SO_4^{2-}(OI) and $SO_4^{2-}{}_{(cb\text{-}unadj)}$, also were similar (see Table 30).

Table 30. Distribution of the Indices of Particulate Air Pollution (in $\mu g/m^3$) in the Reanalysis of the ACS Study

Pollutant[a]	Mean	SD	Percentiles						
			0	5	25	50	75	95	100
$PM_{2.5}$(OI MD)	17.5	5.1	9	10	13	17	21	25	33
$PM_{2.5}$(OI MD) with Denver omitted	17.5	5.1	9	10	13	17	21	25	33
$PM_{2.5}$(DC MD)	17.4	5.3	8	9	13	17	21	25	33
$PM_{2.5}$(DC MD) with Denver omitted	17.6	5.2	9	10	13	17	21	25	33
$PM_{2.5}$(DC)	20.0	5.3	10	12	16	2	23	29	38
PM_{15}(DC)	40.0	9.3	25	29	33	39	4	59	77
$PM_{15-2.5}$(DC)	20.1	7.1	9	11	15	19	23	33	42
PM_{15}(SSI)	58.7	13.0	34	40	51	56	66	84	101
TSP(IPMN)	74.6	16.6	42	50	65	71	85	108	113
TSP	66.4	15.3	40	49	56	64	72	101	127
SO_4^{2-}(DC)	6.7	4.4	0.9	1.9	3.4	6.3	8.9	12.9	27.0
SO_4^{2-}(OI)	10.6	3.6	3.6	4.5	8.1	11.0	13.1	15.7	23.5
$SO_4^{2-}_{(cb\text{-}unadj)}$	10.5	3.4	3.0	4.7	8.0	12.7	12.6	15.7	19.4
$SO_4^{2-}_{(cb\text{-}adj\ US)}$	6.7	3.2	0.0	1.4	4.3	7.0	8.8	11.6	15.0
$SO_4^{2-}_{(cb\text{-}adj\ region)}$	5.9	3.4	0.0	1.0	2.8	6.1	8.1	11.1	17.0
$SO_4^{2-}_{(cb\text{-}adj\ season)}$	6.6	3.1	0.3	1.7	4.2	6.9	8.6	11.7	15.6

[a] Refer to the Abbreviations and Other Terms section at the end of the Investigators' Report for the specific meanings of these pollutant terms and to Table 29 for the sources of pollutant data. All values are means unless indicated by MD (median).

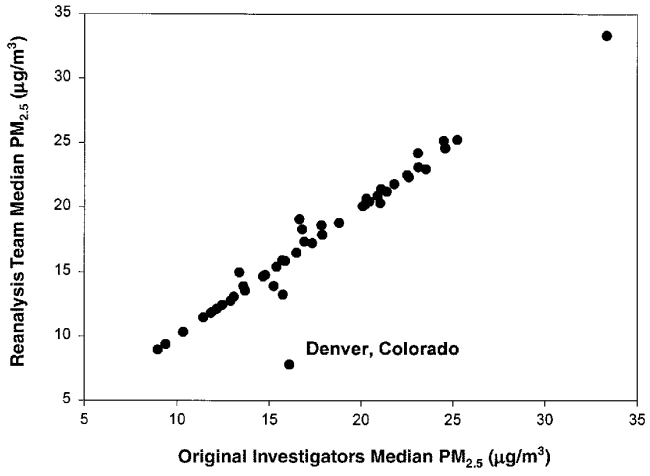

Figure 8. Comparison of fine particle median values between the Original Investigators of the ACS Study and the Reanalysis Team. Reanalysis Team values were based on data from the IPMN.

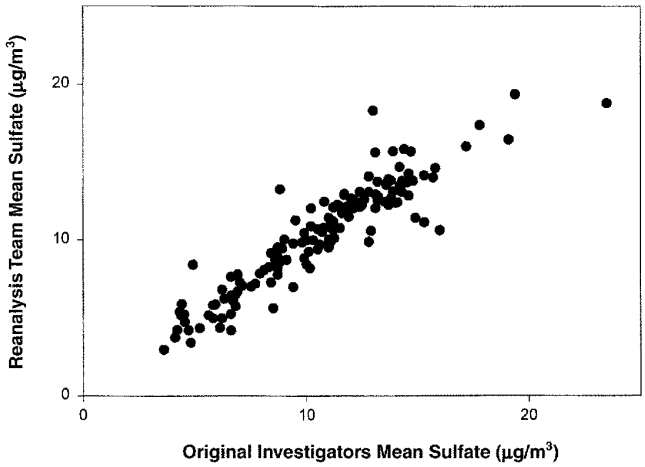

Figure 9. Comparison of mean sulfate values between the Original Investigators of the ACS Study and the Reanalysis Team. Reanalysis Team values were based on 1980 and 1981 data from AIRS and the IPMN.

Table 31. Relative Risks of Mortality from All Causes, Cardiopulmonary Disease, and Lung Cancer Associated with Various Measures of Air Pollution from the Reanalysis of the ACS Study[a]

		Cause of Death		
Pollutant[b]	Number of Cities	All Causes	Cardiopulmonary Disease	Lung Cancer
$PM_{2.5}$(OI MD)	50	1.18 (1.09–1.26)	1.30 (1.17–1.44)	1.00 (0.79–1.28)
$PM_{2.5}$(OI MD) Denver Omitted	49	1.18 (1.10–1.27)	1.30 (1.17–1.44)	0.99 (0.78–1.26)
$PM_{2.5}$(DC MD)	50	1.14 (1.06–1.22)	1.26 (1.14–1.39)	1.08 (0.88–1.32)
$PM_{2.5}$(DC MD) Denver Omitted	49	1.17 (1.09–1.26)	1.28 (1.15–1.42)	1.02 (0.81–1.30)
$PM_{2.5}$(DC)	63	1.12 (1.06–1.19)	1.26 (1.16–1.38)	1.08 (0.88–1.32)
PM_{15}(DC)	63	1.05 (1.01–1.09)	1.09 (1.04–1.15)	1.01 (0.90–1.13)
$PM_{15-2.5}$(DC)	63	1.01 (0.97–1.06)	1.01 (0.95–1.08)	0.97 (0.83–1.13)
PM_{15}(SSI)	59	1.02 (0.99–1.05)	1.07 (1.03–1.11)	0.98 (0.89–1.08)
TSP(IPMN)	58	1.00 (0.98–1.02)	1.02 (0.99–1.05)	0.95 (0.89–1.02)
TSP	156	0.99 (0.98–1.00)	0.99 (0.97–1.01)	0.94 (0.90–0.99)
SO_4^{2-}(DC)	51	1.17 (1.10–1.23)	1.29 (1.19–1.40)	1.09 (0.90–1.33)
SO_4^{2-}(OI)	151	1.15 (1.09–1.21)	1.25 (1.16–1.36)	1.33 (1.10–1.61)
$SO_4^{2-}{}_{(cb\text{-}unadj)}$	144	1.14 (1.07–1.20)	1.24 (1.15–1.35)	1.18 (0.97–1.44)
$SO_4^{2-}{}_{(cb\text{-}adj\ US)}$	144	1.18 (1.11–1.26)	1.31 (1.19–1.43)	1.18 (0.96–1.47)
$SO_4^{2-}{}_{(cb\text{-}adj\ region)}$	144	1.23 (1.16–1.30)	1.34 (1.23–1.45)	1.25 (1.03–1.52)
$SO_4^{2-}{}_{(cb\text{-}adj\ season)}$	144	1.17 (1.09–1.25)	1.29 (1.17–1.42)	1.16 (0.93–1.44)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the ACS Study, this difference for fine particles was 24.5 $\mu g/m^3$, and for sulfate was 19.9 $\mu g/m^3$. Analyses are based on the Extended Model with calendar year as the time axis and the baseline hazard function stratified by 1-year age groups, gender, and race. See Table 19 for a complete list of covariates included in the Extended Model. Data are RRs with 95% CIs.

[b] Refer to the Abbreviations and Other Terms section at the end of the Investigators' Report for the specific meanings of these pollutant terms and to Table 29 for the sources of pollutant data. All values are means unless indicated by MD (median).

For lung cancer mortality, relative risks for fine particles varied around 1.0 with almost all 95% confidence intervals including unity, meaning they were not significant. By comparison, the relative risks for all measures of sulfate were high (from 1.16 to 1.33; except for SO_4^{2-}(DC), which was 1.09), and two of them were statistically significant [SO_4^{2-}(OI) and $SO_4^{2-}{}_{(cb\text{-}adj\ region)}$]. The relative risks of all-cause mortality associated with TSP, TSP(IPMN), PM_{15}(DC), $PM_{15-2.5}$(DC), and PM_{15}(SSI) were less (in the range 0.99 to 1.05) than those for fine particles and sulfate (ranging from 1.12 to 1.23). We observed a similar pattern for cardiopulmonary deaths.

The relative risks of all-cause and cardiopulmonary mortality associated with our estimates of sulfate values for the 144 cities were similar to those that had been obtained by the Original Investigators using their estimates for 151 cities. When we used sulfate values adjusted for the artifactual sulfate with one equation for the entire United States [$SO_4^{2-}{}_{(cb\text{-}adj\ US)}$], we obtained slightly higher relative risks of mortality from all causes and cardiopulmonary disease than we did when we used the unadjusted sulfate concentrations. We calculated these relative risks for a change in sulfate equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; for the unadjusted sulfate data, this value was 19.9 $\mu g/m^3$; for the adjusted sulfate data, this value was 15.0 $\mu g/m^3$. Evaluating the relative risks on the basis of adjusted sulfate values at their corresponding range reduced the size of the effect to that of the unadjusted values. This is not unexpected, as the adjustment is based on a linear equation. The correlation between the adjusted and unadjusted sulfate values was 0.92. The lung cancer risk, however, was much lower (RR = 1.18, 95% CI: 0.96–1.47) if the adjusted sulfate values were employed. This value is somewhat similar to that obtained using sulfate

values from the IPMN on 51 cities (RR = 1.09, 95% CI: 0.90–1.33). Thus the association between sulfate and lung cancer mortality is sensitive to the air pollution data used.

When we evaluated the relative risks of mortality on the basis of adjusted sulfate values for three regions of the United States, or two seasons, the risks were larger than those that were based on a single adjustment for the entire United States. These risks remained higher even if they were evaluated at the respective ranges.

Because several of the cities involved in the ACS Study had limited sulfate data, resulting in potentially unstable estimates of annual averages, we restricted our analysis to those cities with at least 20 observations for sulfate from AIRS. The relative risk of all-cause mortality on the basis of this restricted sample, which included 126 cities, was 1.26 (95% CI: 1.18–1.34). The relative risk that had been calculated from the Original Investigators' sulfate measurements for these same 126 cities was 1.21 (95% CI: 1.14–1.29). These results suggest that there was some instability in risk estimates resulting from city selection (the Original Investigators' risk estimate calculated from measurements in 151 cities was 1.15), but not from the selection of number of observations per city.

We also examined the influence of monitor location on the association between sulfate and mortality by restricting the monitors selected for data analysis to those whose land use code was residential or urban, thereby excluding sites designated as industrial, agricultural, or mobile. This restriction on land use reduced the number of cities available for analysis from 126 to 120 (on the basis of our selection criterion that required at least 20 observations per year). This resulted in only a marginal change in the relative risk of all-cause mortality (RR = 1.24, 95% CI: 1.16–1.32) compared with relative risk calculated using data from the unrestricted 126 cities (RR = 1.26, 95% CI: 1.18–1.34).

Seasonal Effects

The additional air pollution data assembled by the Reanalysis Team permitted an assessment of differences in risk by season. Specifically, we examined the association between the gaseous pollutants (CO, NO_2, O_3, and SO_2) and all-cause, cardiopulmonary, and lung cancer deaths for exposures occurring in two periods of the year: the warmer period of April to September and the cooler period of October to March. We found that sulfur dioxide, nitrogen dioxide, and carbon monoxide concentrations tended to be higher in the cooler time period, whereas ozone levels clearly were elevated during the warmer months (Table 32). For all causes and cardiopulmonary causes of death, the relative risks for each of the four gases examined tended to be higher in the warmer period than in the cooler period. Table 32 shows that the pattern was not as consistent, however, for lung cancer mortality.

FLEXIBLE MODELING

We analyzed the ACS Study data by applying the same flexible spline regression model we used to describe the Six Cities Study data in that Flexible Modeling section. (Further details of this analysis are given in Appendix C.) We used this generalization of the Cox proportional-hazards model to investigate possible nonlinear or time-dependent effects of fine particles and sulfate in the ACS Study. With two exceptions, we used methods similar to those used in the flexible analyses of the Six Cities Study.

First, in our analysis of the Six Cities Study, we had difficulty fully characterizing the shape of the exposure–response

Table 32. Relative Risks of Mortality from All Causes, Cardiopulmonary Disease, and Lung Cancer Associated with Gaseous Copollutants by Season from the Reanalysis of the ACS Study[a]

			Cause of Death		
Pollutant	Season	Seasonal Mean Concentration	All Causes	Cardiopulmonary Disease	Lung Cancer
SO_2 (ppb)	April–September	7.18	1.35 (1.25–1.45)	1.48 (1.33–1.64)	1.40 (1.10–1.79)
	October–March	11.24	1.23 (1.17–1.29)	1.29 (1.20–1.38)	1.00 (0.85–1.18)
NO_2 (ppb)	April–September	23.65	0.96 (0.91–1.02)	0.96 (0.88–1.04)	0.79 (0.65–0.96)
	October–March	27.20	0.93 (0.89–0.97)	0.94 (0.88–0.99)	1.01 (0.87–1.16)
CO (ppm)	April–September	1.33	1.02 (0.96–1.08)	1.00 (0.92–1.09)	0.80 (0.65–0.99)
	October–March	1.73	0.95 (0.90–1.00)	0.90 (0.84–0.97)	0.86 (0.72–1.01)
O_3 (ppb)	April–September	30.44	1.02 (0.96–1.07)	1.08 (1.01–1.16)	0.81 (0.69–0.94)
	October–March	15.07	0.81 (0.76–0.87)	0.82 (0.74–0.91)	0.78 (0.61–0.99)

[a] Analyses based on the Extended Model; see Table 19 for a complete list of covariates included in the Extended Model. Data are RRs with 95% CIs.

curve because observations were available for only six cities. In contrast, the ACS Study included 50 cities in the fine particle cohort and 151 cities in the sulfate cohort, thereby affording a greater opportunity to explore the shape of the exposure–response relation between fine particles and mortality.

The second difference involves the sampling strategy by which we selected random subsets from the ACS cohort, which is much larger than the Six Cities cohort. To conform to the limitations of our flexible modeling software with respect to sample size, we fit the flexible product model to 10 randomly selected disjoint subsets of the ACS Study participants, each including 2,200 individuals. Thus a total of 22,000 individuals, including 1,700 who had died, formed the basis for hypothesis testing using the combined likelihood ratio test discussed in Appendix C. After we accounted for the differences in degrees of freedom, we achieved a level of statistical precision with this combined sample that was comparable to that in the Six Cities Study.

Similar considerations led us to modify our case-cohort approach in order to obtain more stable estimates of the flexible functions of interest. We conducted our modified case-cohort analysis on data from a subset of 2,500 individuals that was created by combining a random subcohort of 1,200 study participants with a random sample of 1,300 deaths.

In our analyses of the 10 random subsets of the ACS cohort, we did not identify a consistent pattern in changes over time on the impact of either fine particles or sulfate on mortality. Whereas the combined likelihood ratio test provided evidence of statistically significant ($P < 0.05$) departures from the Cox proportional-hazards assumption for both fine particles and sulfate, temporal patterns in the hazard ratio varied considerably among subsets, with no pattern being more frequent than any other. The modified case-cohort analyses confirmed the lack of systematic temporal patterns for either fine particles or sulfate; those analyses indicated that the adjusted effects of particles remained nearly constant during the follow-up period (see Figures C.9 and C.13 in Appendix C).

Flexible analyses of the ACS data yielded evidence of nonlinear exposure–response relations ($P < 0.01$) for both fine particles and sulfate. Whereas some differences in subset-specific estimates of the exposure–response relation were apparent, we found evidence of nonlinearity for both fine particles and sulfate. This was confirmed by the case-cohort analysis, which allowed us to estimate the two exposure–response curves more precisely. Figure 10 shows the 3 *df* quadratic spline estimate of the nonlinear effect of fine particles on log–hazard ratio for all-cause mortality, adjusted for pack-years of cigarette smoking for

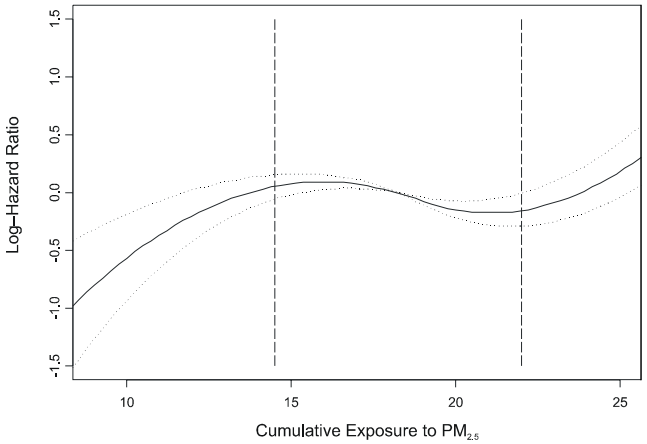

Figure 10. Impact of cumulative exposure to fine particles in the ACS Study. Flexible quadratic spline estimate (3 *df*) of the nonlinear effect of increasing the exposure to fine particles on the log–hazard ratio of mortality in a case-cohort subset of the ACS Study, adjusted for BMI, education level, and pack-years of smoking for current- and former-smokers. The log–hazard ratio was associated with a change in fine particles (24.5 $\mu g/m^3$) equal to the difference in mean concentrations between the most-polluted city and the least-polluted city. Along the horizontal axis, the solid curve represents the point estimate of the log–hazard ratio and the dashed curves the point-wise 95% confidence interval. The left and right dashed vertical lines indicate the first and third quartiles of fine particles in the sample of 2,500 individuals included in the ACS Study.

current- and former-smokers, BMI, and education. This analysis suggests that a monotone exposure–response relation may be limited to the lower half of the range of particle exposures, with little difference in response between moderate and high levels of exposure.

Figure 11 shows the case-cohort estimate of the adjusted effect of sulfate on the log hazard in the ACS Study. The

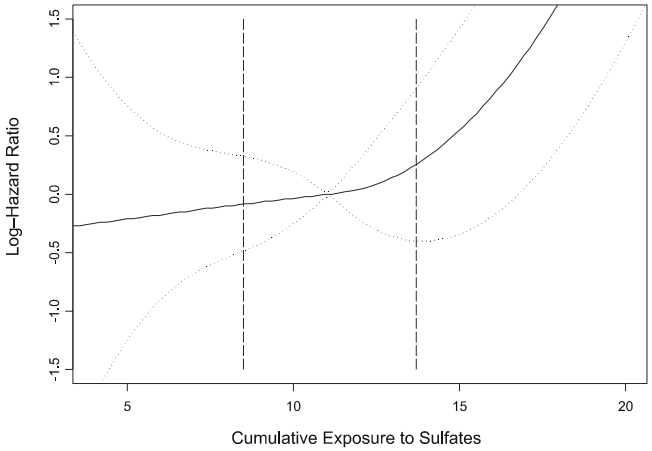

Figure 11. Impact of cumulative exposure to sulfate in the ACS Study. Flexible quadratic spline estimate (3 *df*) of the nonlinear effect of increasing the level of exposure to sulfate on the log–hazard of mortality in a case-cohort subset of the ACS Study. The log–hazard ratio was associated with a change in sulfate (19.9 $\mu g/m^3$) equal to the difference in mean concentrations between the most-polluted city and the least-polluted city. The solid curve represents the point estimate of the log–hazard ratio and the dashed curves represent the point-wise 95% condidence interval.

exposure–response curve is nonlinear even though it is monotone. The curve is quite flat in the lower half of the range of observed sulfate levels, corresponding to exposures below about 14 $\mu g/m^3$. At higher exposures, however, sulfate is associated with a relatively sharp increase in mortality. This pattern is consistent with that observed in most of the random subsets (see Figure C.12 in Appendix C).

Although the curve in Figure 11 depicts only the relative impact of changes in sulfate on mortality, we can calculate the magnitude of this impact by multiplying the estimate in Figure 11 by the time-dependent estimate (refer to Appendix C). When we combine the two estimates, our results show that the impact of sulfate on mortality is quite modest; the hazard ratio that corresponds to a change from the minimum to the maximum of the 151 city-specific sulfate levels does not exceed 1.10, which is comparable to the results that had been obtained by the Original Investigators in their analysis of the ACS data.

Using flexible analyses of the ACS data, we also demonstrated a nonlinear relation between BMI and mortality. Figure 12 shows the 3 *df* estimate of the adjusted effect of BMI, based on the case-cohort approach. The relation between BMI and log hazard is U-shaped, as it was in the analysis of the Six Cities Study data, with risk increasing at both low and high values of BMI. This suggests that the association between BMI and mortality may be approximated well by a quadratic function.

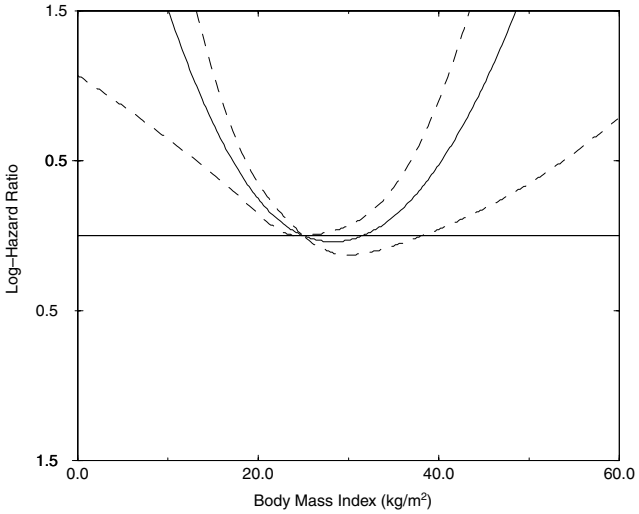

Figure 12. Flexible nonlinear estimate of the effect of BMI in the ACS Study. Flexible quadratic spline (3 *df*) estimate of the nonlinear effect of increasing body mass index on the log–hazard of mortality in a case-cohort subset of the ACS Study. The log–hazard ratio is plotted with respect to the mean BMI as reference value. The solid curve represents the point estimates of the log–hazard ratio and the dashed curves represent the pointwise 95% confidence interval.

ECOLOGIC COVARIATES

Both the Six Cities Study and the ACS Study had included a number of covariates in the risk models that had been developed by the Original Investigators, in addition to the main covariate of interest, namely, ambient particle levels. Individual-level data (data supplied by subjects via questionnaires) were available for each covariate included in the models, except fine particle air pollution, which was measured at the city level. Because particle levels were represented at the ecologic rather than the individual level, it is conceivable that the associations that had been observed between particles (particularly fine particles and sulfate) and mortality in these two studies could have been due at least in part to other city-level characteristics correlated with both air pollution and mortality. To assess the possibility of such ecologic confounding, the Reanalysis Team obtained data on a number of ecologic covariates not considered by the Original Investigators, and examined the effect of including these new city-level covariates on the air pollution–mortality association. Because the Six Cities Study had involved, at most, 5 *df* for incorporating additional ecologic covariates, we restricted our analysis to the ACS Study, which involved 151 cities in the sulfate cohort and 50 cities in the fine particle cohort.

Selection of Ecologic Covariates

The Reanalysis Team applied several criteria in selecting the additional ecologic covariates to be included in this component of the sensitivity analyses. First, a potential ecologic covariate had to represent a valid measure of a well-defined attribute of each city. Second, there had to be a plausible biological or social mechanism by which a candidate ecologic covariate could affect mortality. And third, we required access to reliable data on those ecologic covariates selected for inclusion in the reanalysis.

There are essentially three related types of ecologic variables. "Aggregated" ecologic variables are derived by aggregating characteristics that have been measured at the individual level to obtain a city-level summary measure. Such aggregated variables are often used as surrogates for measures of individual-level variables. In other words, most aggregated variables are considered to have direct analogs at the individual level. "Group-level" variables represent attributes that have individual analogs, but usually are obtained from measurements at the city level (eg, maximum daily exposure to ozone). "Global" or "contextual" variables refer to attributes of cities that do not have analogs at the individual level (eg, total area of green space, or population density). Although contextual variables represent group-level attributes, they may be

aggregated from individual data. For example, income disparity as measured by the Gini coefficient, which must be calculated from individual income data, has no analog at the individual level (Kaplan et al 1996). Other relevant contextual variables include city-level unemployment or poverty, both of which measure constructs other than individual employment or poverty status.

Although aggregated ecologic variables are sometimes used as substitutes for measurements of the same constructs at the individual level, the results of such analyses often do not represent the level of association that would be measured had individual-level variables been used. This is referred to as ecologic (or cross-level) bias, and has been discussed extensively in the literature (eg, Piantadosi et al 1988; Greenland and Morgenstern 1989; Brenner et al 1992 a,b; Greenland and Robbins 1994).

The ACS Study can be considered a hybrid design, in that detailed individual information had been collected, but the primary exposure variables (fine particles and sulfate) had been derived from measurements taken at the city level. Thus, although the study had provided data on the joint distribution of many covariates across the study population, exposure had been assessed on an ecologic level. This may not be a serious difficulty, depending on the extent to which city-level ambient air pollution concentrations, as estimated from regularly collected data at the beginning of the study period, had represented the relevant exposure metric for individuals. However, if these exposure metrics also had represented certain characteristics inherent to the city, and these were correlated with other city-level characteristics, then it is possible that there could have been residual confounding (on the ecologic level) by these other city-level characteristics. Our purpose in this set of sensitivity analyses was to select plausible ecologic covariates that we could use to address this last concern. We thus attempted to identify variables that, rather than mimicking individual traits, measured essential characteristics of the cities (ie, contextual variables).

In selecting ecologic covariates for this component of the sensitivity analyses, the Reanalysis Team drew upon the literature on the determinants of population health. Evans and Stoddart (1990) outlined a number of contextual risk factors for health, including the social environment, the physical environment, and health care. These three categories guided our search for possible ecologic confounders (Figure 13). We searched Medline to find evidence of links between mortality and specific contextual variables within these categories. For those variables for which the literature indicated a possible health risk, we sought data that had been collected in the early 1980s on counties and metropolitan areas from US government sources. Although data were not available at this level of geographic resolution for all potential contextual variables, we did identify a number of relevant ecologic covariates for inclusion in the reanalysis.

A detailed description of the process we used to select those ecologic covariates included in our sensitivity analyses of the ACS Study is given in Appendix E (which is available on request from the Health Effects Institute). In order to ensure that the ecologic covariate values were representative of the MSAs included in the ACS Study, we carefully examined the geographic area spanned by all 158 MSAs that had been considered by the Original Investigators (details are provided in Appendix F, which is available on request from the Health Effects Institute). The city-specific values of the ecologic covariates we selected are listed in Appendix G (also available on request from the Health Effects Institute).

As described in Appendix E, we selected 20 ecologic covariates suitable to include in the reanalysis from a longer list of 30 potential variables (Table 33). Eight measures of the social environment were considered: population change, percentage of white residents, percentage of black residents, mean income of residents in 1979, poverty level in 1979, income disparity as measured by the Gini

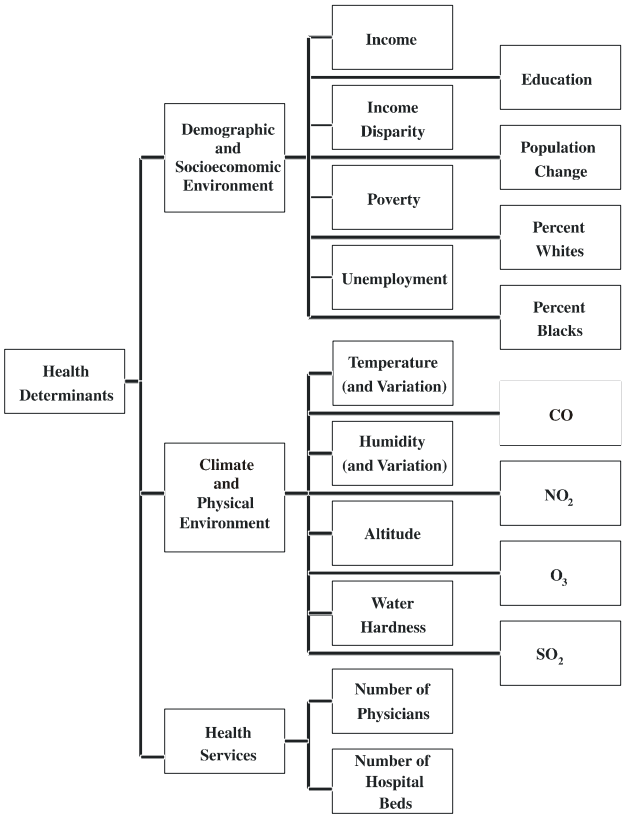

Figure 13. Summary of selected ecologic covariates.

Table 33. A Summary of the Ecologic Covariates and the Sources of Data Used in the Reanalysis of the ACS Study

Ecologic Covariate	Number of Cities: Sulfate	Number of Cities: Fine Particles	Description of Covariate and Source of Data
Demographic Factors			
Population change	139	48	Percentage of net change in number of residents between 1980 and 1986; US Bureau of the Census, 1986 Population Estimates by County with Components of Change[a]
Whites	151	50	Percentage of persons residing in the MSA in 1980 who classified themselves as being of white race; US Bureau of the Census, County Population Estimates (experimental) by Age, Sex, and Race: 1980–1984[a]
Blacks	151	50	Percentage of persons residing in the MSA in 1980 who classified themselves as being of black race; 1980 Census of Population and Housing (April 1, 1980), STF3 data[b]
Socioeconomic Factors			
Income	151	50	Mean annual per capita income in US dollars for 1979; 1980 Census of Population and Housing (April 1, 1980), STF3 data[b]
Poverty	151	50	Percentage of individuals in 1979 who were classified as living below the poverty level specific to their family size, age, and number of dependents; US Bureau of the Census, Current Population Reports, Series P-26, Numbers 86-NE-SC, 86-S-SC, 86-ENC-SC, 86-WNC-SC, and 86-W-SC; and 1980 Census of Population and Housing, STF3 data[a]
Income disparity	151	50	Gini coefficient (see Selection of Ecologic Covariates section for description) calculated from income group data for 1979 as outlined in Shyrock et al 1976; 1980 Census of Population and Housing (April 1, 1980), STF3 data[b]
Unemployment	151	50	Percentage of total civilian labor force who were unemployed in 1986; US Bureau of Labor Statistics, Employment and Unemployment in States and Local Areas, Annual, 1986[a]
Education	151	50	Percentage of the number of persons 25 years of age or older who indicated they had completed 4 years of high school or some years of college divided by the total number of persons 25 years and older; 1980 Census of Population and Housing (April 1, 1980), STF3 data[b]
Health Services			
Physicians	138	48	Number of professionally active, non-Federal physicians with known addresses per 100,000 residents as of July 1, 1985; American Medical Association's Physician Characteristics and Distribution in the US, 1986[a]
Hospital beds	139	48	Number of hospital beds per 100,000 residents as of July 1, 1985; survey (September 30, 1985) of all hospitals (registered and unregistered) excluding old-age homes, convalescent homes, and sanatoriums; American Hospital Association's Hospital Statistics, 1986[a]
Climate			
Temperature	135	46	Maximum daily temperature (°F) averaged by month for 1980 through 1989; the average of all monthly averages was used as the ecologic covariate; data provided directly to us by the US National Climatic Data Center of the National Oceanic and Atmospheric Administration (NOAA), Asheville NC
Temperature variation	135	46	Variation in maximum daily temperature (°F) averaged by month for 1980 through 1989; the average of the monthly variation was used as the ecologic covariate; data provided directly to us by NOAA
Relative humidity	95	37	Minimum daily relative humidity (%) averaged by month for 1984 through 1989; the mean of all monthly averages was used as the ecologic covariate; data provided directly to us by NOAA
Relative humidity variation	95	37	Variation in minimum daily relative humidity (%) averaged by month for 1984 through 1989; the average of the monthly variation was used as the ecologic covariate; data provided directly to us by NOAA
Physical Environment			
Altitude	110	38	Measured as meters above sea level; US Places (24000+); from Environmental Systems Research Institute (1999)
Water hardness	109	49	Concentration of $CaCO_3$ (ppm) in drinking water, measured ca 1970; National Institutes of Health data cited in Feinleib et al 1979
Gaseous Copollutants			
CO	107	44	Daily average concentrations averaged by year for 1980; from residential, commercial, or mobile monitors
NO_2	74	33	Daily average concentrations averaged by year for 1980; from residential, commercial, or mobile monitors
O_3	117	45	Daily 1-hour maximum concentrations
SO_2	113	38	Daily average concentrations averaged by year for 1980; from residential, commercial, or mobile monitors

[a] Cited in the County and City Data Book (1988).

[b] Data from Geolytics Software (1999).

coefficient, unemployment in 1986, and percentage of residents age 25 or older who had completed high school. We obtained two measures of the provision of health care services: number of physicians per 100,000 residents and number of hospital beds per 100,000 residents. In terms of the physical environment, we considered altitude, water hardness, and climate (average maximum temperature, average monthly variation in maximum temperature, average daily relative humidity, and average monthly variation in daily relative humidity). We used four gaseous copollutants in these analyses as well: carbon monoxide, nitrogen dioxide, ozone, and sulfur dioxide.

Unfortunately, we were unable to obtain data on certain ecologic covariates for some of the cities included in the ACS Study. In particular, data on relative humidity and the gaseous copollutants were sparse. For this reason, we sometimes conducted the sensitivity analyses that used those ecologic covariates on subsets of the original cohort of cities. The numbers of cities for which we obtained values for the selected ecologic covariates in both the sulfate and fine particle cohort of the ACS Study are given in Table 33.

Incorporation of Ecologic Covariates in Cox Regression

The Reanalysis Team examined the effect of these ecologic covariates on the association between particulate air pollution and mortality by incorporating them in the Cox proportional-hazards regression model in the same way that air pollution, itself an ecologic covariate, had been used in the Cox models employed by the Original Investigators. Instead of using the Original Model as the basis of our analyses, however, we used the Extended Model developed in the Alternative Risk Models section for the ACS Study data. This permitted us to make full use of all the individual-level covariates as well as examining the effects of these additional ecologic covariates.

Table 34 summarizes the Extended Model results of including, in turn, each of the 20 ecologic covariates selected by the Reanalysis Team. The first column of data in this table shows the relative risks of all-cause mortality associated with sulfate exposure in a model without the ecologic covariate. Note that because values for some of the ecologic covariates were not available for some cities, the relative risks of mortality associated with sulfate vary somewhat depending on the number of cities for which the ecologic covariate data were available (see Table 33). The second column of data shows the relative risk of all-cause mortality associated with sulfate exposure, with the ecologic covariate included in the Extended Model. The inclusion of most of these ecologic covariates does not appear to have a marked impact on the relative risk of all-cause mortality for sulfate. The inclusion of population change, which has an inverse association with mortality (RR = 0.85, 95% CI: 0.81–0.89) and is correlated negatively with sulfate ($r = -0.40$), decreases the relative risk from 1.15 to 1.06 in the Extended Model, reducing the excess relative risk from 0.15 to 0.06. The inclusion of sulfur dioxide, which has a positive association with mortality (RR = 1.30, 95% CI: 1.23–1.38) and is positively correlated with sulfate ($r = 0.48$), reduces the relative risk from 1.16 to 1.04. The lower confidence limits on the relative risk adjusted for sulfur dioxide (RR = 1.04, 95% CI: 0.98–1.11) is less than 1.00, resulting in the loss of formal statistical significance after adjustment.

Adjustment for the effects of ecologic covariates in this manner requires careful interpretation. Abramowicz and colleagues (2000) have shown that the inclusion of ecologic covariates in the model results in a downward bias, unlike the case of linear regression, in the estimated relative risk of the exposure of primary interest (fine particle air pollution in the present case). Although this bias may be small, some degree of bias toward the null value of unity appears to persist regardless of the strength of the association between the ecologic covariate and mortality, or between the ecologic covariate and particle levels.

The relative risks and associated confidence limits for the ecologic covariates themselves are shown in the last two columns in Table 34. The relative risk of mortality associated with population change in the Extended Model excluding sulfate is 0.85, with a 95% CI (0.81–0.89) that excluded the null value of 1.00. (Inclusion of sulfate in the model increases the relative risk of population change only slightly, from 0.85 to 0.87.) Because population change is thus a strong ecologic predictor of all-cause mortality, the reduction in the relative risk of sulfate (from 1.15 to 1.06) could be an overadjustment. However, the extent of overadjustment is difficult to judge without further information on the nature of the relation between population change and mortality.

Other covariates that appear to be significantly associated with mortality in the absence of sulfate include hospital beds (RR = 1.13, at the range of the values among the cities for which such health services data were available), income (RR = 0.93), income disparity as measured by the Gini coefficient (RR = 0.88), unemployment (RR = 1.12), temperature (RR = 0.88), temperature variation (RR = 1.18), and water hardness (RR = 1.08). With the exception of sulfur dioxide, none of the gaseous copollutants (CO, NO_2, and O_3) demonstrated a positive and clearly significant association with mortality.

The Reanalysis Team also employed three multivariate models that permitted simultaneous adjustment for more than one ecologic covariate. The first multivariate model

Table 34. Relative Risks of Mortality from All Causes Associated with an Increase in Sulfate After Adjusting for Selected Ecologic Covariates[a]

Ecologic Covariate	Relative Risk from Sulfate		Relative Risk from Ecologic Covariate	
	Without Ecologic Covariate	With Ecologic Covariate	Without Sulfate	With Sulfate
Demographic Factors				
Population change	1.15 (1.08–1.21)	1.06 (0.99–1.13)	0.85 (0.81–0.89)	0.87 (0.82–0.91)
Whites	1.15 (1.09–1.21)	1.18 (1.11–1.25)	1.02 (0.98–1.06)	1.06 (1.02–1.11)
Blacks	1.15 (1.09–1.21)	1.17 (1.10–1.24)	1.01 (0.96–1.06)	0.96 (0.91–1.01)
Socioeconomic Factors				
Income	1.15 (1.09–1.21)	1.15 (1.08–1.21)	0.93 (0.88–0.97)	0.93 (0.88–0.97)
Poverty	1.15 (1.09–1.21)	1.15 (1.09–1.22)	0.95 (0.91–1.00)	0.94 (0.90–0.99)
Income disparity	1.15 (1.09–1.21)	1.15 (1.09–1.21)	0.88 (0.84–0.93)	0.88 (0.83–0.93)
Unemployment	1.15 (1.09–1.21)	1.13 (1.06–1.19)	1.12 (1.06–1.19)	1.09 (1.03–1.16)
Education	1.15 (1.09–1.21)	1.13 (1.06–1.20)	0.91 (0.86–0.96)	0.96 (0.90–1.02)
Health Services				
Physicians	1.15 (1.08–1.21)	1.14 (1.08–1.21)	0.95 (0.89–1.01)	0.96 (0.90–1.01)
Hospital beds	1.15 (1.08–1.21)	1.13 (1.07–1.20)	1.13 (1.06–1.21)	1.12 (1.04–1.19)
Climate				
Temperature	1.14 (1.07–1.21)	1.11 (1.04–1.17)	0.88 (0.85–0.92)	0.90 (0.86–0.94)
Temperature variation	1.14 (1.07–1.21)	1.10 (1.04–1.17)	1.18 (1.11–1.24)	1.16 (1.09–1.22)
Relative humidity	1.13 (1.05–1.21)	1.14 (1.05–1.24)	1.05 (0.99–1.12)	0.98 (0.91–1.06)
Relative humidity variation	1.13 (1.05–1.21)	1.16 (1.08–1.25)	0.96 (0.90–1.02)	0.92 (0.86–0.98)
Physical Environment				
Altitude	1.10 (1.04–1.18)	1.16 (1.08–1.24)	1.05 (0.99–1.12)	1.12 (1.04–1.19)
Water hardness	1.11 (1.04–1.18)	1.12 (1.05–1.19)	1.08 (1.02–1.13)	1.08 (1.03–1.14)
Gaseous Copollutants				
CO	1.14 (1.08–1.21)	1.14 (1.08–1.21)	0.98 (0.92–1.03)	0.97 (0.92–1.03)
NO_2	1.11 (1.04–1.19)	1.14 (1.06–1.22)	0.93 (0.89–0.98)	0.91 (0.87–0.96)
O_3	1.15 (1.08–1.22)	1.15 (1.09–1.22)	0.93 (0.87–0.99)	0.92 (0.86–0.98)
SO_2	1.16 (1.09–1.23)	1.04 (0.98–1.11)	1.30 (1.23–1.38)	1.28 (1.20–1.37)
Multiple Covariate Analyses				
Gaseous copollutants	1.09 (1.02–1.17)	1.00 (0.93–1.09)		1.36 (1.26–1.46)[b]
Socioeconomic status[c]	1.15 (1.08–1.21)	1.10 (1.02–1.18)		
25%[d]	1.10 (1.02–1.18)	0.99 (0.90–1.09)		1.25 (1.14–1.38)[b]

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the ACS Study, this difference for sulfate was 19.9 $\mu g/m^3$. Analyses are based on the Extended Model with calendar year as the time axis and the baseline hazard function stratified by 1-year age groups, gender, and race. See Table 19 for a complete list of covariates included in the Extended Model. Data are RRs with 95% CIs.

[b] Whenever SO_2 was included in a multivariate model, the relative risk of mortality associated with it is given in the far right column.

[c] Includes population change and the five socioeconomic factors (income, poverty, income disparity, unemployment, and education).

[d] Incorporated all ecologic covariates that, when analyzed individually in a bivariate model, were found to produce a change of 25% or more in the excess relative risk associated with the pollutant of interest. In this analysis: population change, temperature variation, altitude, and SO_2.

included all socioeconomic variables as well as population change. The second model included all four gaseous copollutants (CO, NO_2, O_3, and SO_2). The third model included all those ecologic covariates that, when analyzed individually in a bivariate model, had resulted in a change of 25% or more in the excess relative risk of mortality associated with the pollutant of interest. Whereas the first ecologic multivariate regression model was intended to provide maximal adjustment for socioeconomic determinants of mortality, the second was designed to isolate the effects of sulfate from gaseous copollutants. The third model, which sought to identify potential confounders empirically, was based on a strategy similar to that employed by Gérin and colleagues (1998); they used logistic regression analysis in a large-scale population-based case-control study of cancer incidence in relation to nearly 400 chemical and other agents found in the workplace.

Some caution is required in the interpretation of the relative risks for sulfate under the multivariate Cox regression models, both because of the possibility of overadjustment noted previously, and because of the moderately high correlation among some of the ecologic covariates considered here (see Appendix G). For example, the correlation between poverty rate and mean income in the sulfate cohort is –0.58 (Table G.7, Appendix G). Similarly, the correlation between mean maximum daily temperature and income disparity is +0.60, because income disparity tends to be greater in the southern United States. The highest correlation (r = –0.96) occurs between percent white and percent black, although none of our multivariate models included both population subgroups.

The relative risk of sulfate alone (RR = 1.15) is reduced following simultaneous adjustment for population change and the five socioeconomic factors (RR = 1.10), although the adjustment is less than that obtained with population change alone (RR = 1.06). This reduced adjustment, which incorporates population change in the multivariate model, is attributable to the complex structure of correlation among the variables in this model. On the other hand, simultaneous adjustment for all four gaseous copollutants leads to a relative risk of sulfate (RR = 1.00) that is less than that following adjustment for sulfur dioxide alone (RR = 1.04). This is similar to the relative risk obtained (RR = 0.99) after adjustment for those covariates (population change, variation in maximum temperature, altitude, and sulfur dioxide) that induced a 25% change in the relative risk from sulfate alone. Again, note that there is a possibility of overadjustment in these latter two cases.

The results of incorporating these same ecologic covariates in the Extended Model for cardiopulmonary and lung cancer mortality are shown in Tables 35 and 36, respectively. A number of ecologic covariates (population change, income, income disparity, unemployment, education, physicians, hospital beds, temperature variation, relative humidity, water hardness, and sulfur dioxide) appear to be associated with cardiopulmonary mortality (Table 35), although etiologic hypotheses underlying these associations are not readily apparent in all cases. Nonetheless, adjustment for these ecologic covariates does not alter the original conclusions concerning the positive association between cardiopulmonary mortality and sulfate exposure. Most of the ecologic covariates do not appear to have a marked impact on estimated relative risks of lung cancer mortality for sulfate (Table 36), although adjustment for altitude reduces the relative risk from 1.24 to 1.14. In the Extended Model runs that ex-cluded sulfate, relative humidity, altitude, and ozone all appear to be associated with lung cancer mortality.

Similar ecologic analyses were carried out for the fine particle cohort. The relative risk of all-cause mortality for fine particles, as with sulfate, was diminished after adjustment for population change or sulfur dioxide (Table 37). This same effect was observed for cardiopulmonary mortality (Table 38). Because lung cancer mortality was not associated with fine particles, no adjustment for ecologic covariates was attempted in this case.

To a certain extent, the effects of the ecologic covariates on the relative risks of mortality for fine particle air pollution can be explained by spatial convergence between the ecologic covariate and exposure to fine particles. Fine particles and sulfate were both highest in the Ohio Valley and around Gary IN, and decreased slightly toward the South and more dramatically toward the West. The two ecologic covariates correlated most highly with both fine particles and sulfate were population change and education. For all-cause mortality, the percentage of the population with high school education or more is inversely associated with mortality from all causes (Tables 34 and 37) and from cardiopulmonary disease (Tables 35 and 38). Our results also bear out the observation that areas where the population has increased tend to have lower mortality.

Although we can postulate a possible biological relation between mortality and exposure to fine particles and sulfate, we cannot suggest that population change is a cause of death. Rather, population change is considered to be an indicator of the economic climate of a metropolitan area. A health effect is associated with the economic climate of a place. Healthy people tend to migrate out of areas of recession to areas experiencing economic well-being, whereas unhealthy people tend to stay where they are. Moreover, areas of heavy manufacturing, which are likely to have

Table 35. Relative Risks of Mortality from Cardiopulmonary Disease Associated with an Increase in Sulfate After Adjusting for Selected Ecologic Covariates[a]

Ecologic Covariate	Relative Risk from Sulfate		Relative Risk from Ecologic Covariate	
	Without Ecologic Covariate	With Ecologic Covariate	Without Sulfate	With Sulfate
Demographic Factors				
Population change	1.24 (1.15–1.35)	1.12 (1.03–1.23)	0.80 (0.75–0.85)	0.83 (0.77–0.90)
Whites	1.25 (1.16–1.36)	1.30 (1.20–1.42)	1.02 (0.97–1.08)	1.09 (1.03–1.16)
Blacks	1.25 (1.16–1.36)	1.28 (1.17–1.40)	1.03 (0.96–1.10)	0.95 (0.88–1.03)
Socioeconomic Factors				
Income	1.25 (1.16–1.36)	1.25 (1.16–1.36)	0.87 (0.81–0.93)	0.87 (0.81–0.94)
Poverty	1.25 (1.16–1.36)	1.26 (1.16–1.36)	0.99 (0.93–1.05)	0.97 (0.91–1.04)
Income disparity	1.25 (1.16–1.36)	1.25 (1.16–1.36)	0.90 (0.83–0.97)	0.90 (0.83–0.97)
Unemployment	1.25 (1.16–1.36)	1.20 (1.11–1.30)	1.27 (1.16–1.38)	1.21 (1.11–1.32)
Education	1.25 (1.16–1.36)	1.21 (1.11–1.32)	0.84 (0.77–0.92)	0.92 (0.83–1.00)
Health Services				
Physicians	1.24 (1.15–1.35)	1.23 (1.14–1.34)	0.86 (0.79–0.94)	0.87 (0.80–0.95)
Hospital beds	1.24 (1.15–1.35)	1.23 (1.13–1.33)	1.18 (1.07–1.30)	1.15 (1.04–1.26)
Climate				
Temperature	1.22 (1.13–1.33)	1.19 (1.09–1.29)	0.87 (0.46–1.64)	0.89 (0.83–0.95)
Temperature variation	1.22 (1.13–1.33)	1.17 (1.08–1.28)	1.26 (1.16–1.36)	1.22 (1.12–1.32)
Relative humidity	1.21 (1.09–1.34)	1.20 (1.06–1.35)	1.11 (1.01–1.21)	1.01 (0.91–1.13)
Relative humidity variation	1.21 (1.09–1.34)	1.28 (1.15–1.42)	0.93 (0.85–1.01)	0.86 (0.79–0.95)
Physical Environment				
Altitude	1.20 (1.10–1.32)	1.26 (1.14–1.39)	1.02 (0.93–1.12)	1.12 (1.02–1.23)
Water hardness	1.19 (1.09–1.30)	1.20 (1.09–1.31)	1.11 (1.03–1.20)	1.12 (1.04–1.21)
Gaseous Copollutants				
CO	1.27 (1.17–1.38)	1.27 (1.17–1.38)	0.94 (0.87–1.02)	0.94 (0.86–1.01)
NO_2	1.25 (1.14–1.38)	1.29 (1.17–1.42)	0.93 (0.87–1.01)	0.89 (0.83–0.97)
O_3	1.26 (1.16–1.37)	1.27 (1.17–1.38)	0.98 (0.90–1.08)	0.96 (0.88–1.05)
SO_2	1.28 (1.18–1.40)	1.14 (1.04–1.25)	1.40 (1.29–1.52)	1.33 (1.22–1.46)
Multiple Covariate Analyses				
Gaseous copollutants	1.22 (1.10–1.35)	1.11 (1.00–1.25)		1.41 (1.27–1.57)[b]
Socioeconomic status[c]	1.24 (1.15–1.35)	1.18 (1.06–1.31)		
25%[d]	1.20 (1.06–1.36)	1.17 (0.97–1.41)		1.38 (1.17–1.61)[b]

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the ACS Study, this difference for sulfate was 19.9 $\mu g/m^3$. Analyses are based on the Extended Model with calendar year as the time axis and the baseline hazard function stratified by 1-year age groups, gender, and race. See Table 19 for a complete list of covariates included in the Extended Model. Data are RRs with 95% CIs.

[b] Whenever SO_2 was included in a multivariate model, the relative risk of mortality associated with it is given in the far right column.

[c] Includes population change and the five socioeconomic factors (income, poverty, income disparity, unemployment, and education).

[d] Incorporated all ecologic covariates that, when analyzed individually in a bivariate model, were found to produce a change of 25% or more in the excess relative risk associated with the pollutant of interest. In this analysis: population change, relative humidity variation, altitude, and SO_2.

Table 36. Relative Risks of Mortality from Lung Cancer Associated with an Increase in Sulfate After Adjusting for Selected Ecologic Covariates[a]

	Relative Risk from Sulfate		Relative Risk from Ecologic Covariate	
Ecologic Covariate	Without Ecologic Covariate	With Ecologic Covariate	Without Sulfate	With Sulfate
Demographic Factors				
Population change	1.31 (1.08–1.58)	1.30 (1.05–1.61)	0.91 (0.78–1.06)	1.00 (0.84–1.18)
Whites	1.33 (1.10–1.61)	1.34 (1.10–1.64)	0.96 (0.84–1.10)	1.03 (0.89–1.19)
Blacks	1.33 (1.10–1.61)	1.35 (1.10–1.66)	1.05 (0.89–1.23)	0.95 (0.80–1.13)
Socioeconomic Factors				
Income	1.33 (1.10–1.61)	1.33 (1.10–1.61)	1.01 (0.86–1.19)	1.02 (0.86–1.20)
Poverty	1.33 (1.10–1.61)	1.33 (1.10–1.61)	0.97 (0.83–1.13)	0.96 (0.82–1.12)
Income disparity	1.33 (1.10–1.61)	1.32 (1.09–1.60)	0.88 (0.73–1.07)	0.89 (0.73–1.08)
Unemployment	1.33 (1.10–1.61)	1.31 (1.08–1.60)	1.13 (0.92–1.38)	1.05 (0.86–1.29)
Education	1.33 (1.10–1.61)	1.34 (1.09–1.65)	0.90 (0.74–1.10)	1.02 (0.82–1.27)
Health Services				
Physicians	1.31 (1.08–1.58)	1.30 (1.07–1.57)	0.91 (0.74–1.11)	0.93 (0.76–1.13)
Hospital beds	1.31 (1.08–1.58)	1.32 (1.09–1.60)	0.92 (0.73–1.16)	0.89 (0.71–1.13)
Climate				
Temperature	1.37 (1.12–1.67)	1.36 (1.11–1.67)	0.94 (0.81–1.09)	0.98 (0.84–1.15)
Temperature variation	1.37 (1.12–1.67)	1.39 (1.13–1.71)	1.00 (0.83–1.21)	0.94 (0.77–1.14)
Relative humidity	1.53 (1.20–1.95)	1.39 (1.04–1.86)	1.37 (1.10–1.72)	1.17 (0.90–1.53)
Relative humidity variation	1.53 (1.20–1.95)	1.49 (1.16–1.92)	1.19 (0.96–1.47)	1.09 (0.88–1.37)
Physical Environment				
Altitude	1.24 (1.00–1.54)	1.14 (0.91–1.44)	0.72 (0.56–0.93)	0.76 (0.58–0.99)
Water hardness	1.31 (1.06–1.62)	1.31 (1.05–1.62)	0.94 (0.78–1.12)	0.94 (0.79–1.13)
Gaseous Copollutants				
CO	1.26 (1.03–1.54)	1.26 (1.03–1.54)	0.82 (0.68–1.00)	0.82 (0.68–1.00)
NO_2	1.28 (1.02–1.62)	1.33 (1.05–1.69)	0.91 (0.76–1.09)	0.87 (0.73–1.05)
O_3	1.27 (1.05–1.55)	1.30 (1.07–1.59)	0.74 (0.59–0.92)	0.72 (0.58–0.90)
SO_2	1.31 (1.07–1.62)	1.36 (1.08–1.72)	1.06 (0.87–1.30)	0.93 (0.74–1.17)
Multiple Covariate Analyses				
Gaseous copollutants	1.31 (1.08–1.58)	1.20 (0.94–1.54)		0.86 (0.66–1.13)[b]
Socioeconomic status[c]	1.42 (1.11–1.82)	1.61 (1.21–2.14)		
25%[d]	1.53 (1.20–1.95)	1.39 (1.04–1.86)		

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the ACS Study, this difference for sulfate was 19.9 $\mu g/m^3$. Analyses are based on the Extended Model with calendar year as the time axis and the baseline hazard function stratified by 1-year age groups, gender, and race. See Table 19 for a complete list of covariates included in the Extended Model. Data are RRs with 95% CIs.

[b] Whenever SO_2 was included in a multivariate model, the relative risk of mortality associated with it is given in the far right column.

[c] Includes population change and the five socioeconomic factors (income, poverty, income disparity, unemployment, and education).

[d] Incorporated all ecologic covariates that, when analyzed individually in a bivariate model, were found to produce a change of 25% or more in the excess relative risk associated with the pollutant of interest. In this analysis: relative humidity.

Table 37. Relative Risks of Mortality from All Causes Associated with an Increase in Fine Particles After Adjusting for Selected Ecologic Covariates[a]

	Relative Risk from Fine Particles		Relative Risk from Ecologic Covariate	
Ecologic Covariate	Without Ecologic Covariate	With Ecologic Covariate	Without Fine Particles	With Fine Particles
Demographic Factors				
Population change	1.19 (1.10–1.28)	1.07 (0.99–1.17)	0.84 (0.80–0.89)	0.86 (0.81–0.92)
Whites	1.18 (1.09–1.26)	1.28 (1.18–1.38)	1.04 (0.98–1.10)	1.14 (1.07–1.22)
Blacks	1.18 (1.09–1.26)	1.25 (1.15–1.37)	1.00 (0.94–1.08)	0.89 (0.82–0.97)
Socioeconomic Factors				
Income	1.18 (1.09–1.26)	1.18 (1.09–1.26)	0.93 (0.87–0.99)	0.93 (0.87–0.99)
Poverty	1.18 (1.09–1.26)	1.22 (1.13–1.32)	0.91 (0.85–0.99)	0.86 (0.79–0.93)
Income disparity	1.18 (1.09–1.26)	1.23 (1.14–1.32)	0.85 (0.80–0.90)	0.83 (0.78–0.88)
Unemployment	1.18 (1.09–1.26)	1.16 (1.08–1.25)	1.09 (1.02–1.16)	1.06 (0.99–1.13)
Education	1.18 (1.09–1.26)	1.17 (1.08–1.27)	0.94 (0.89–1.00)	1.00 (0.93–1.06)
Health Services				
Physicians	1.19 (1.10–1.28)	1.19 (1.10–1.28)	0.95 (0.88–1.04)	0.95 (0.87–1.03)
Hospital beds	1.19 (1.10–1.28)	1.18 (1.10–1.27)	1.05 (0.96–1.16)	1.03 (0.93–1.14)
Climate				
Temperature	1.13 (1.05–1.22)	1.12 (1.03–1.20)	0.86 (0.81–0.90)	0.86 (0.81–0.91)
Temperature variation	1.13 (1.05–1.22)	1.07 (0.99–1.16)	1.18 (1.11–1.25)	1.16 (1.09–1.23)
Relative humidity	1.19 (1.08–1.30)	1.19 (1.08–1.31)	1.03 (0.96–1.10)	1.00 (0.93–1.07)
Relative humidity variation	1.19 (1.08–1.30)	1.22 (1.11–1.34)	0.96 (0.90–1.04)	0.93 (0.86–1.00)
Physical Environment				
Altitude	1.11 (1.02–1.20)	1.14 (1.05–1.24)	1.06 (1.00–1.13)	1.09 (1.02–1.16)
Water hardness	1.17 (1.09–1.26)	1.16 (1.07–1.25)	1.15 (1.08–1.23)	1.14 (1.07–1.21)
Gaseous Copollutants				
CO	1.17 (1.09–1.26)	1.18 (1.10–1.27)	0.93 (0.88–0.98)	0.92 (0.87–0.97)
NO_2	1.15 (1.05–1.25)	1.22 (1.11–1.33)	0.95 (0.89–1.01)	0.90 (0.84–0.96)
O_3	1.16 (1.08–1.25)	1.18 (1.10–1.27)	0.90 (0.84–0.97)	0.88 (0.82–0.95)
SO_2	1.20 (1.11–1.29)	1.03 (0.95–1.13)	1.49 (1.36–1.64)	1.46 (1.32–1.63)
Multiple Covariate Analyses				
Gaseous copollutants	1.15 (1.05–1.25)	1.07 (0.96–1.19)		1.46 (1.31–1.63)[b]
Socioeconomic status[c]	1.19 (1.10–1.28)	1.15 (1.04–1.28)		
25%[d]	1.09 (0.99–1.20)	1.33 (1.09–1.61)		1.14 (0.90–1.45)[b]

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the ACS Study, this difference for fine particles was 24.5 $\mu g/m^3$. Analyses are based on the Extended Model with calendar year as the time axis and the baseline hazard function stratified by 1-year age groups, gender, and race. See Table 19 for a complete list of covariates included in the Extended Model. Data are RRs with 95% CIs.

[b] Whenever SO_2 was included in a multivariate model, the relative risk of mortality associated with it is given in the far right column.

[c] Includes population change and the five socioeconomic factors (income, poverty, income disparity, unemployment, and education).

[d] Incorporated all ecologic covariates that, when analyzed individually in a bivariate model, were found to produce a change of 25% or more in the excess relative risk associated with the pollutant of interest. In this analysis: population change, whites, poverty, income disparity, temperature, altitude, NO_2, and SO_2.

Table 38. Relative Risks of Mortality from Cardiopulmonary Disease Associated with an Increase in Fine Particles After Adjusting for Selected Ecologic Covariates[a]

Ecologic Covariate	Relative Risk from Fine Particles		Relative Risk from Ecologic Covariate	
	Without Ecologic Covariate	With Ecologic Covariate	Without Fine Particles	With Fine Particles
Demographic Factors				
Population change	1.29 (1.16–1.44)	1.12 (0.99–1.27)	0.78 (0.72–0.84)	0.81 (0.74–0.89)
Whites	1.30 (1.17–1.44)	1.43 (1.27–1.61)	1.02 (0.94–1.11)	1.17 (1.07–1.29)
Blacks	1.30 (1.17–1.44)	1.39 (1.23–1.58)	1.04 (0.95–1.15)	0.88 (0.78–0.99)
Socioeconomic Factors				
Income	1.30 (1.17–1.44)	1.30 (1.17–1.44)	0.90 (0.82–0.99)	0.90 (0.82–0.99)
Poverty	1.30 (1.17–1.44)	1.34 (1.20–1.49)	0.95 (0.85–1.06)	0.87 (0.78–0.98)
Income disparity	1.30 (1.17–1.44)	1.35 (1.22–1.50)	0.86 (0.79–0.93)	0.82 (0.76–0.89)
Unemployment	1.30 (1.17–1.44)	1.25 (1.12–1.39)	1.19 (1.09–1.30)	1.13 (1.03–1.24)
Education	1.30 (1.17–1.44)	1.27 (1.13–1.42)	0.89 (0.82–0.96)	0.96 (0.88–1.06)
Health Services				
Physicians	1.29 (1.16–1.44)	1.30 (1.17–1.44)	0.88 (0.78–1.00)	0.88 (0.78–0.99)
Hospital beds	1.29 (1.16–1.44)	1.28 (1.15–1.43)	1.16 (1.02–1.33)	1.13 (0.99–1.30)
Climate				
Temperature	1.21 (1.09–1.35)	1.19 (1.07–1.33)	0.83 (0.77–0.90)	0.84 (0.77–0.91)
Temperature variation	1.21 (1.09–1.35)	1.13 (1.01–1.27)	1.24 (1.14–1.35)	1.21 (1.10–1.32)
Relative humidity	1.22 (1.07–1.39)	1.20 (1.04–1.37)	1.09 (0.99–1.21)	1.06 (0.96–1.17)
Relative humidity variation	1.22 (1.07–1.39)	1.27 (1.10–1.45)	0.94 (0.85–1.05)	0.90 (0.81–1.00)
Physical Environment				
Altitude	1.23 (1.09–1.38)	1.25 (1.11–1.41)	1.02 (0.93–1.11)	1.06 (0.97–1.16)
Water hardness	1.28 (1.15–1.43)	1.26 (1.13–1.41)	1.21 (1.10–1.32)	1.19 (1.09–1.31)
Gaseous Copollutants				
CO	1.30 (1.17–1.45)	1.32 (1.19–1.47)	0.92 (0.85–1.00)	0.74 (0.61–0.90)
NO_2	1.32 (1.16–1.49)	1.39 (1.22–1.59)	0.99 (0.91–1.08)	0.91 (0.83–1.00)
O_3	1.29 (1.16–1.44)	1.30 (1.17–1.45)	0.96 (0.86–1.07)	0.93 (0.83–1.04)
SO_2	1.35 (1.21–1.51)	1.17 (1.03–1.33)	1.59 (1.39–1.81)	1.45 (1.25–1.69)
Multiple Covariate Analyses				
Gaseous copollutants	1.32 (1.16–1.49)	1.22 (1.05–1.43)		1.42 (1.21–1.65)[b]
Socioeconomic status[c]	1.29 (1.16–1.44)	1.17 (1.01–1.36)		
25%[d]	1.26 (1.12–1.42)	1.19 (1.00–1.41)		1.26 (1.04–1.52)[b]

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the ACS Study, this difference for fine particles was 24.5 $\mu g/m^3$. Analyses are based on the Extended Model with calendar year as the time axis and the baseline hazard function stratified by 1-year age groups, gender, and race. See Table 19 for a complete list of covariates included in the Extended Model. Data are RRs with 95% CIs.

[b] Whenever SO_2 was included in a multivariate model, the relative risk of mortality associated with it is given in the far right column.

[c] Includes population change and the five socioeconomic factors (income, poverty, income disparity, unemployment, and education).

[d] Incorporated all ecologic covariates that, when analyzed individually in a bivariate model, were found to produce a change of 25% or more in the excess relative risk associated with the pollutant of interest. In this analysis: population change, whites, temperature variation, and SO_2.

high levels of both fine particles and sulfate, are also likely to have experienced a recession during the 1980s.

Adding ecologic covariates to the Cox proportional-hazards regression models provides one method of controlling for ecologic confounding. However, this method does not enable researchers to control for the spatial autocorrelation that can result from missing or unmeasured ecologic covariates. Moreover, statistical tests of significance are reliable only when researchers can be sure that the residuals of their models are not autocorrelated. For these reasons, the Reanalysis Team used spatial smoothing and filtering techniques (described in the following section) to characterize the spatial patterns in the data and to model the data in a way that makes explicit provision for spatial autocorrelation.

SPATIAL ANALYSES

An important issue in the analysis of data from the ACS Study is whether or not the observations are independent or correlated. Residents of cities located near one another may be at similar risk of mortality resulting from shared aspects of their social and physical environments, such as socioeconomic influences, access to health care, dietary habits, and environmental or occupational exposures, which lead to spatial autocorrelation in the data. Other covariates used in both the original analyses and the present sensitivity analyses may be spatially correlated as well. Spatial autocorrelation affects the statistical power of associations, with positive spatial autocorrelation in the residuals increasing the likelihood of a false-positive finding. Failure to control for spatial autocorrelation can invalidate traditional tests of significance and lead to biased estimates of coefficients in regression models. Recent studies have shown that when correlation is taken into account, formerly significant findings in ordinary-least-squares regression analyses may become insignificant (Griffith et al 1998).

For the purposes of the ACS data analyses described in the previous sections, the Reanalysis Team assumed the statistical independence of all observations. This assumption was necessary for the application of the standard Cox proportional-hazards regression model of survival that had been used by the Original Investigators. In this section, we consider analytic techniques that allow for the possibility of spatial autocorrelation in the ACS data.

Within the context of the general spatial analysis framework shown in Figure 14, the Reanalysis Team explored different approaches for capturing spatial patterns in the data. The specific approaches included in this framework and their associated results are discussed in later sections. We begin with a description of spatial patterns in the data and the methods used to summarize them in smooth maps of the variables of interest; then we discuss formal tests for spatial autocorrelation. Computationally simple two-stage random effects regression methods, used to take into account spatial clustering at both city and broader regional levels, provide risk estimates that are in close agreement with estimates derived from a new random effects Cox regression model specifically developed by the Reanalysis Team. We also employed spatial filtering methods to remove broad spatial patterns in the data before we applied our two-stage regression models. Further details of the spatial smoothing and spatial filtering methods used by the Reanalysis Team are given in Appendix H, which is available on request from Health Effects Institute.

Spatial Analysis Framework

Under the Cox proportional-hazards regression model of survival used by the Original Investigators, it had been assumed that the survival of each individual could be represented by statistically independent random variables. However, several processes involved in predicting mortality in space and time may induce some degree of statistical dependence in the data. We have attempted to characterize some of these processes and model the corresponding statistical dependence within the context of the spatial analysis framework shown in Figure 14.

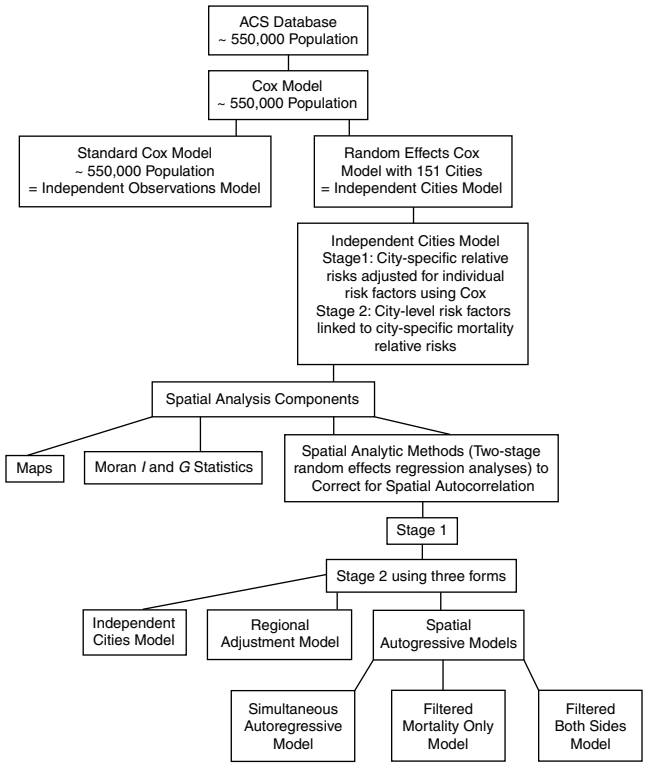

Figure 14. Paradigm of spatial analyses.

We discuss in detail in Appendix H the rationale for taking spatial dependence into account. When epidemiologic investigations use health data from contiguous or nearby geographic areas, the data may not provide independent estimates of the dependent variable (in this case, relative risk of mortality). If we account for this lack of independence with covariates that are also spatially autocorrelated, then no statistical problem arises because the error terms from such a model tend to be uncorrelated. However, if areas differ in some unmeasured or unsuspected way that affects mortality, residuals are likely to be correlated (Cook and Pocock 1983). Autocorrelated errors can result in overestimates of significance. Therefore, to ensure reliable significance tests, we needed to account for spatial dependence in the regression models. This could also lead to the identification of new covariates that may explain some of the variation in mortality that manifests as autocorrelation in the residuals. Careful examination and mapping of the residuals can suggest locations where the model fails to predict mortality accurately, thereby providing useful information on omitted covariates.

Following the lead of the Original Investigators, we began by applying the Cox proportional-hazards regression model of survival to individual-level data on members of the sulfate cohort (approximately 550,000 subjects) or the fine particle cohort (approximately 295,000 subjects) from the ACS Study data. This standard Cox model includes a number of risk factors measured at the individual level, with the baseline hazard function stratified by age, gender, and race. Because a key assumption in this model is that the survival times for all subjects are statistically independent, this approach is referred to as the Independent Observations Model.

Next we extended the standard Cox model to allow the baseline hazard function to vary at random among cities, resulting in a random effects Cox model (Appendix I, available on request from the Health Effects Institute). The random effects Cox model is predicated on the assumption that there is more variation in individual risks of mortality than can be explained by the standard Cox model, and that subjects in the same city are expected to have more similar mortality risks than subjects living in different cities. This modeling framework also implies that the city itself is a risk factor for mortality in the sense that even after we control for all available risk factor information at the individual and ecologic level, the city in which a subject lives will have some influence on his or her survival.

Our random effects Cox model assumes that differences in the mortality risks of individuals in different cities are independent of the proximity of those cities, so that mortality is spatially clustered beyond that associated with specific cities. To date, we have not extended the random effects Cox model to incorporate such spatial dependence, although such an extension appears to be technically feasible. Instead, in a two-stage approach, we have exploited spatial regression methods developed for normally distributed data (Getis and Ord 1996) to address the extra spatial variation in mortality beyond that induced by clustering at the city level.

In the first stage, we fitted a standard Cox model to the individual-level data, including an indicator variable for each city. (With this approach, one city must be selected as an index in order to compare mortality between each city and the index city.) No city-level or ecologic variables, including air pollution, were included at this stage. Because estimates of the city-specific mortality rates relative to the index city are, by definition, correlated, we transformed the covariance matrix of the estimated mortality rates in each city relative to the index city to independence using methods developed by Easton and colleagues (1991). We then linked the logarithms of the comparative city-specific relative risks to the ecologic covariates, including air pollution, using a linear model with independent errors of the form $\tau^2 + v_j$, in which τ^2 is the unexplained variation in the true logarithms of the city-specific comparative mortality rates and v_j is the uncertainty in the estimated mortality rate for city j obtained in the first stage. We arrived at estimates of τ^2 by the method of moments (DerSimonian and Laird 1986), and used weighted least squares to estimate the effects of the ecologic covariates and air pollution, with weights given by $1/(\tau^2 + v_j)$ for the jth city.

Because both the random effects and two-stage approaches arrive at mortality rates for each city that are assumed to be independent, they are referred to as Independent Cities Models. Although their estimates of the city-level covariate effects and τ^2 are expected to be similar, these estimates will not be equivalent because of the nonlinear structure of the random effects regression model. Under the assumption that there is no extra variation as a result of clustering of mortality rates within a city (ie, $\tau^2 = 0$), the parameter estimates of the covariate effects and their uncertainty under the random effects and the standard Cox models will, however, be identical.

We then questioned the assumption that the cities' mortality rates are independent. Even after we have controlled for available risk factor information at both the individual and ecologic level, the adjusted city-specific risk estimates conceivably could exhibit evidence of spatial autocorrelation. Such spatial dependence in the adjusted risk estimates could result from unidentified processes that vary in space and lead to clustering of mortality rates. As discussed

in Appendix H, the Moran *I* statistic can be used to test for the presence of spatial autocorrelation in the city-specific mortality rates.

We addressed the statistical form of such spatial dependence using four approaches. In the first approach (Regional Adjustment Model) we removed spatial variation in mortality rates by adjusting the city-specific values for broad regional patterns. Specifically, we used the seven regions of the United States (Figure 15) defined in the National Morbidity and Mortality Air Pollution Study conducted by Samet and colleagues (2000), and removed differences in mortality rates between regions together with effects of fine particle air pollution and other variables measured at the city level. Using arbitrary regions, however, provides only limited and ad hoc control for spatial dependence because such regions have not been derived on the basis of prior knowledge or empirical evidence about what processes may have caused the spatial dependence. Thus, it is possible that such regions may overcontrol or undercontrol the actual spatial dependence in the model. As the regions adopted for this analysis were quite large, smaller-area spatial dependence was probably neglected.

Our second spatial analysis approach involved modeling the broader spatial patterns of mortality in city-specific regions using the Moran *G* statistic, which is designed to detect local autocorrelation that arises when variables display nonstationarity (spatial dependence is not the same everywhere in the United States, but varies locally over space). In this approach, the values for relative risk of mortality that surround a city within a given distance are divided by their global average values for the entire United States. Cities that are surrounded by other cities with elevated mortality rates will have significantly elevated values. Then a *G* statistic is calculated for every city in the dataset. Next, the mortality rate at that city is multiplied by the expected value of the city's *G* statistic, and then divided by the actual value. This has the effect of removing spatial dependence from the mortality rate or other covariate of interest, including air pollution. A summary *G* statistic was calculated by averaging the city-specific *G* statisics from all cities.

Two criteria were used to select the distance for defining the regions. First, a graphic technique known as semivariogram analysis assessed the distance at which spatial dependence among mortality, air pollution, and other covariates diminished. Second, we selected the distance that minimized the residual spatial autocorrelation in the variable, as determined by Moran *I,* a global test of spatial autocorrelation. Iterations between these two analyses suggested that regions with a radius of approximately 600 km would remove spatial dependence from the variables without inducing negative autocorrelation (in which, for example, high mortality associates with neighboring low mortality).

This approach is similar to that for time-series mortality studies in which adjustments for temporal trends in mortality rates employ multiday moving-average windows. Similar to time-series models, negative autocorrelation suggests that the filtering procedure has removed too much information from the variable, including some of the attribute value that is not specifically associated with spatial arrangement. The spatially filtered city-specific mortality rates were then regressed on the unfiltered sulfate concentrations (this technique is referred to as the Filtered Mortality Model).

The third modeling approach involved spatially filtering not only the city-specific relative risks of mortality but also the sulfate concentrations and other covariates (referred to as Filtered Both Sides Model, or the Spatial Filtering Model). Here, a 600-km radius also was sufficient to remove any evidence of spatial autocorrelation in the sulfate data. In this approach, we compared mortality rates and sulfate levels after removing broad spatial patterns in both datasets. Clustering of mortality rates by city is taken into account in the filtering approach using weighted least squares regression with weights given by $1/(\tau^2 + v_j)$.

This filtering approach has a number of advantages over the ad hoc regional adjustment. First, because every city is assigned its own region, this method provides a more sensitive adjustment to spatial dependence. Unlike the Regional Adjustment Model, filtering methods rely on actual measured spatial dependence in the data. Second, because it is based on the *G* statistic, the filter deals explicitly with nonstationarities in the variables. Third, this method relies on a linear regression model that is easily interpreted, unlike the simultaneous autoregression technique discussed below. Fourth, the selection of the filter distance forces the analyst to think carefully about what may have caused the spatial dependence.

We found that fairly broad regional patterns with a radius of approximately 600 km appeared to exert a major influence on both mortality and pollution. As mentioned above, it is possible to overfilter the data (ie, by removing not only the spatial pattern, but also some of the attribute value not associated with spatial arrangement) used in the deterministic part of the model, thus removing part of the possible causal relation between air pollution and mortality. One can minimize this problem by carefully selecting the filter distance such that little or no negative autocorrelation is induced in the filtered variables or in the residuals.

Another potential weakness of this approach lies in the binary structure of the spatial weighting matrix. All cities

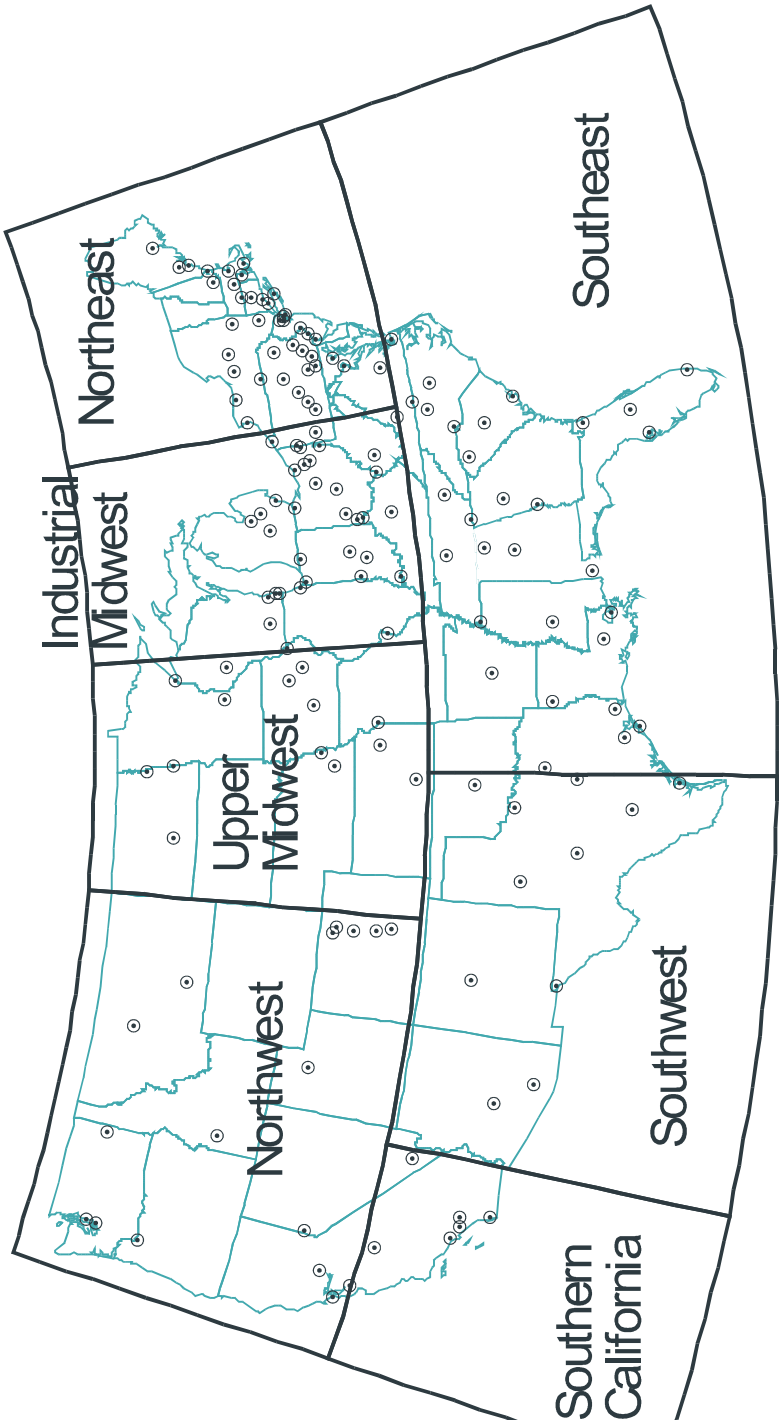

Figure 15. Regions of the US used for the Regional Adjustment Model. Cities in the ACS sulfate cohort are noted (Y). The regions were designed for use in Samet and associates (2000).

within the 600-km radius are assigned an equal weight of one, whereas those outside this distance are weighted zero. With many spatial patterns, we would expect to see a distance decay effect, whereby spatial dependence diminishes as a function of distance away from the city of interest. Refinements of the weighting matrix to account for distance decay were not possible in this analysis, but this represents an important area for future research.

We further examined the robustness of our results to the method of controlling for spatial autocorrelation by using a fourth spatial analysis approach referred to as the Simultaneous Autoregressive Model. In this approach, the logarithms of the city-specific mortality rates are the response variables, assumed to be normally distributed, and the city-level or ecologic covariates are used as predictors. The error structure in the Simultaneous Autoregressive Model explicitly incorporates correlation among mortality rates after accounting for city-level predictors of mortality. The correlation structure is predicated on the nearest-neighbor concept, which assumes that one is more likely to be influenced by one's neighbor, no matter how far away that neighbor is, than by one that is not a neighbor.

We defined a city's neighbors in the following manner. First, Thiessen polygons (geographic areas that incorporate all points closer to the given city than to any other) were constructed for each city. The neighbors of any city were then defined as the area enclosed by all the Thiessen polygons touching the polygon of that city. Each city may have a different number of neighbors, and the nearest neighbor may be a different distance away for each city. The correlation structure then derived correlates a city's residual response only with the residual responses of its neighbors; cities that are not neighbors were assumed to be uncorrelated. We assumed a common correlation parameter for the entire dataset, which was estimated simultaneously with the regression parameters using maximum likelihood techniques in S-PLUS statistical software (MathSoft, Seattle WA). We also weighted the analysis by $1/(\tau^2 + v_j)$, thus incorporating the concept of a random effects model in the analysis.

We also considered an adjusted nearest-neighbor approach, in which mortality rates were assumed to be correlated among cities when the cities were nearest neighbors or were within the average distance between cities (111 km for cities with sulfate data and 123 km for cities with sulfur dioxide data). Here we report only the results for the nearest-neighbor approach, because the results obtained using the adjusted nearest-neighbor approach were virtually identical. (Although the data used to generate the correlation matrix using the adjusted nearest-neighbor approach incorporated more cities in the Northeast and Ohio Valley regions than did the data that used nearest neighbors only, the inclusion of these additional cities did not markedly influence the estimate of the common correlation parameter.)

Both approaches we used in the Simultaneous Autoregressive Model (Thiessen polygons and adjusted nearest neighbor) relied on a more localized spatial dependence assumption (the nearest neighbor) than did the Regional Adjustment Model or spatial filtered models. We suggest that spatial interaction among risk factors is most likely to occur among neighbors, regardless of distance. For most places, except parts of the Western United States where the sampled coverage of cities is sparse, this provided a more localized control on spatial autocorrelation. With this type of model, we did not try to understand the mechanisms underlying autocorrelation. Simultaneous estimation of the coefficients and the autoregressive component of the error term did not allow for intensive investigation into spatial relations. Our intention instead was to incorporate, within the error structure of the model, a term that accounts for autocorrelation so that the remaining errors are uncorrelated and therefore amenable to reliable significance testing.

Mortality rates and pollution display an east-to-west trend, whereby the values of both are generally higher in the east than in the west. When spatial relations are not the same in all directions, we refer to the spatial pattern in the data as being anisotropic. In large datasets, it is possible to build in allowance for this directionality. We attempted to remove this trend with a special form of regression known as a trend surface. A trend surface includes the actual geographic coordinates of the cities as independent predictor variables. Both the Simultaneous Autoregressive Model and the Regional Adjustment Model rely implicitly on the assumption that no trend is present (isotropy) in the data. The results of these models must be viewed with this limitation in mind. The filtering procedure, although it implicitly assumes isotropy in the radius around a city, at least removes all autocorrelation from the data. Semivariogram analyses suggested that the filter also removed the east-to-west trend from the data. Thus, the spatial filtering approach has the advantage of producing reliable estimates even when anisotropy is present. Incorporation of this directional trend into future analyses may improve the robustness of the results.

Using each of these four spatial analysis approaches (Regional Adjustment Model, Filtered Mortality Model, Spatial Filtering [Filtered Both Sides] Model, and Simultaneous Autoregressive [Nearest Neighbor] Model) affords an opportunity to examine the extent to which inferences about the association between ambient air pollution and

mortality are influenced by spatial patterns in the data. Under the Independent Observations Model we employed originally, we assumed that all observations on individual cohort members were statistically independent. Our spatial analysis paradigm goes well beyond the Independent Observations Model by allowing for spatial patterns in the data. Specifically, spatial clustering is considered at the city level in the Independent Cities Model, at the regional level in the Regional Adjustment Model, and at the broader spatial level in the Spatial Filtering Model and the Simultaneous Autoregressive (Nearest Neighbor) Model.

An important practical aspect of our work was the use of two-stage regression models to address spatial autocorrelation. This two-stage approach was validated first in the Independent Observations Model by comparing the results with those from the standard Cox model, and then in the Independent Cities Model by comparing the results with those from the random effects Cox model. This validation of the two-stage regression approach supports its use in the more complex spatial filtering models, for which more direct approaches are not yet available.

The different approaches included in our framework of spatial analyses may be viewed as affording greater levels of control for spatial autocorrelation. The Independent Observations Model, which assumes that no spatial autocorrelation exists, represents a baseline with which the results of the spatial analytic techniques may be compared. The Independent Cities Model takes into account clustering of mortality rates by city, but does not acknowledge spatial autocorrelation at a broader regional level. The Regional Adjustment Model does for spatial dependence, but only within seven predetermined regions of the United States. The Spatial Filtering Model allows for more general spatial patterns either in relative risks or in both air pollution and mortality (the filtered both sides model). By filtering out broad spatial patterns in the data, these latter models seek to associate local variations in mortality rates (adjusted for regional mortality rates) with air pollution.

Comparison of the risk of mortality associated with air pollution estimated by the different spatial analytic techniques can suggest whether spatial association exists on the broader regional scale or the narrower subregional scale. We discuss the application of these spatial analytic methods to the ACS data, beginning with visual evidence of spatial patterns in the data.

Spatial Patterns in the Data

Spatial patterns in the data can be assessed by visualization, exploration, and modeling (Bailey and Gatrell 1995). Graphic visualization is an important first step toward understanding spatial patterns in the variables of interest. We then can use exploratory spatial methods to examine spatial concordance in key variables such as mortality rates and indices of air pollution, along with other ecologic covariates we assembled for the cities included in the ACS Study. After this initial examination, we can use spatial modeling techniques to assess and describe spatial autocorrelation, and to develop spatial regression methods describing any association between the covariates of interest and mortality.

The spatial distributions of sulfate, sulfur dioxide, and fine particles are displayed graphically in Figures 16, 17, and 18, respectively. We derived the smoothed surfaces of pollutant concentrations by using kriging methods to interpolate values between the cities for which direct measurements of these pollutants were available. In these three maps, darker colors represent higher pollution levels. High levels of both sulfate and sulfur dioxide tend to cluster in the Lower Great Lakes area. They exhibit similar spatial patterns, although sulfur dioxide is spatially more concentrated than sulfate. This observation raises the possibility that some of the effect attributed to sulfate pollution in the original ACS Study may have resulted from sulfur dioxide. High values for fine particles are clustered slightly farther south. All three pollutants exhibit higher concentrations in the east than in the west.

The uncertainty in these kriged estimates is considered in Appendix H (see Figures H.3, H.4, and H.5). In general, uncertainty is larger in those areas where the point samples (cities) are less dense. The standard errors of the kriged estimates are largest in the Upper Midwest where the point coverage is most sparse. The standard errors of most of the interpolated sulfur dioxide levels are less than 1.1 ppb, corresponding to a pointwise 95% confidence limit of ± 2.2 ppb or less; in most of the Eastern United States, the confidence limits are within ± 1.37 ppb. For sulfate, the largest errors are also in the Upper Midwest and along the borders of Oregon and Wyoming. In these areas, the contour surfaces are accurate to within ± 1.17 $\mu g/m^3$. For large portions of the Eastern United States, the errors are less than ± 0.39 $\mu g/m^3$.

Uncertainties in the fine particle dataset are somewhat larger because there are fewer cities in the fine particle cohort (50) than in the sulfate cohort (151). The Upper Midwest again displays the largest errors, with estimates of fine particle levels accurate to within ± 2.1 $\mu g/m^3$. The most precise estimates are found for the Lower Great Lakes area and the Southeast, where the monitoring networks are densest. For most of this region, predictions are accurate to within ± 1.37 $\mu g/m^3$.

Spatial overlays of the mortality rates and air quality levels are shown in Figures 19, 20, and 21. We prepared

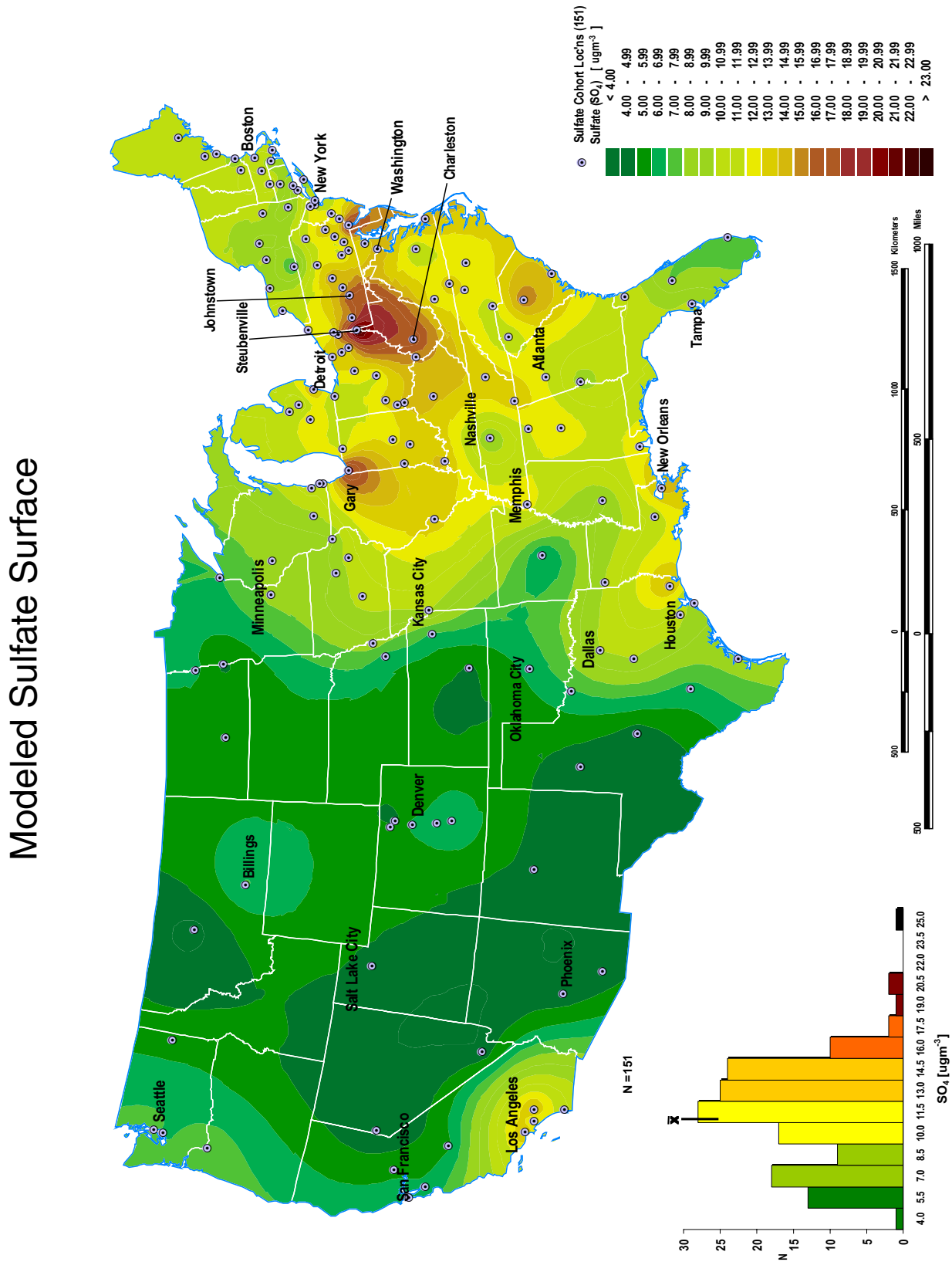

Figure 16. Spatial distribution of sulfate.

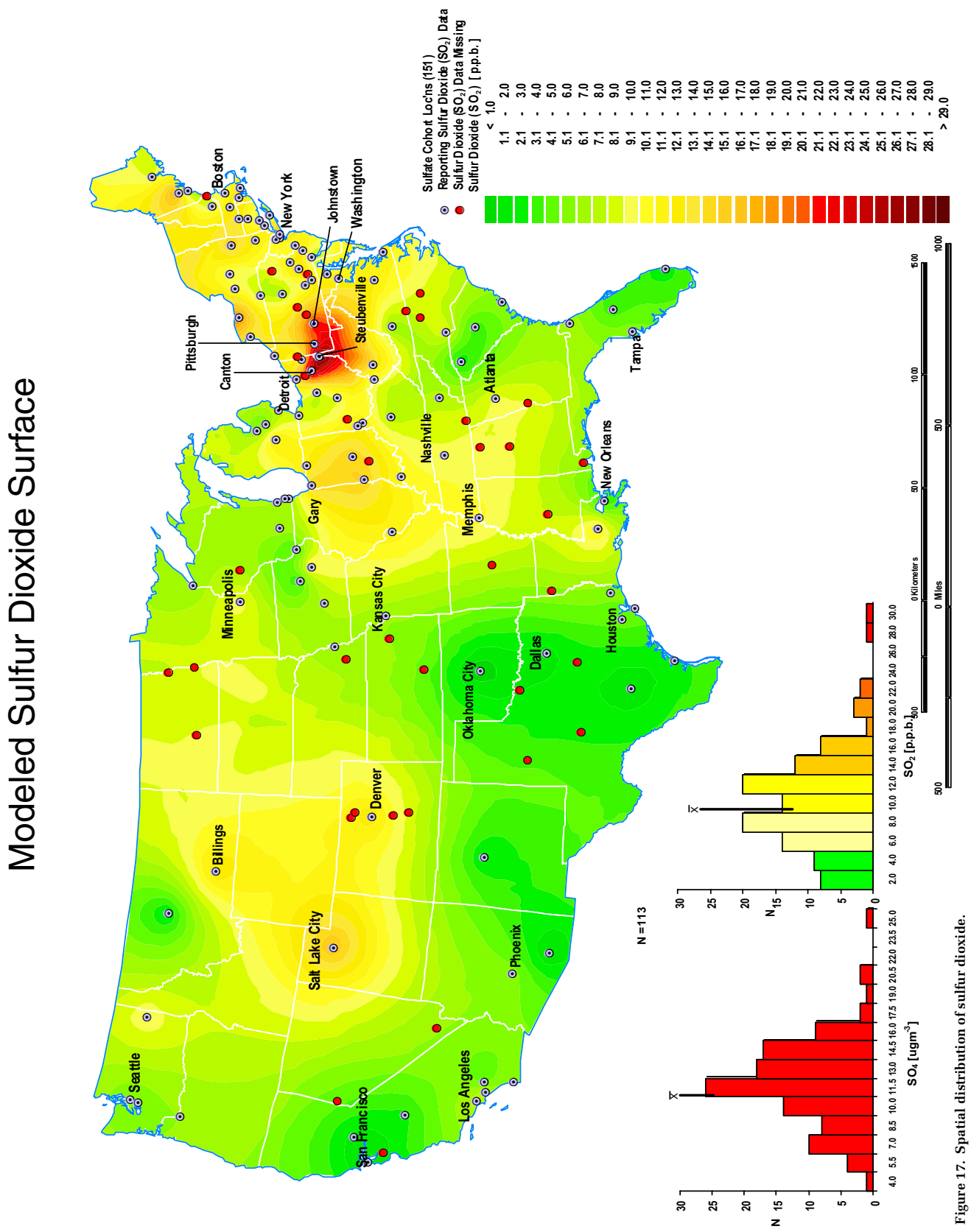

Figure 17. Spatial distribution of sulfur dioxide.

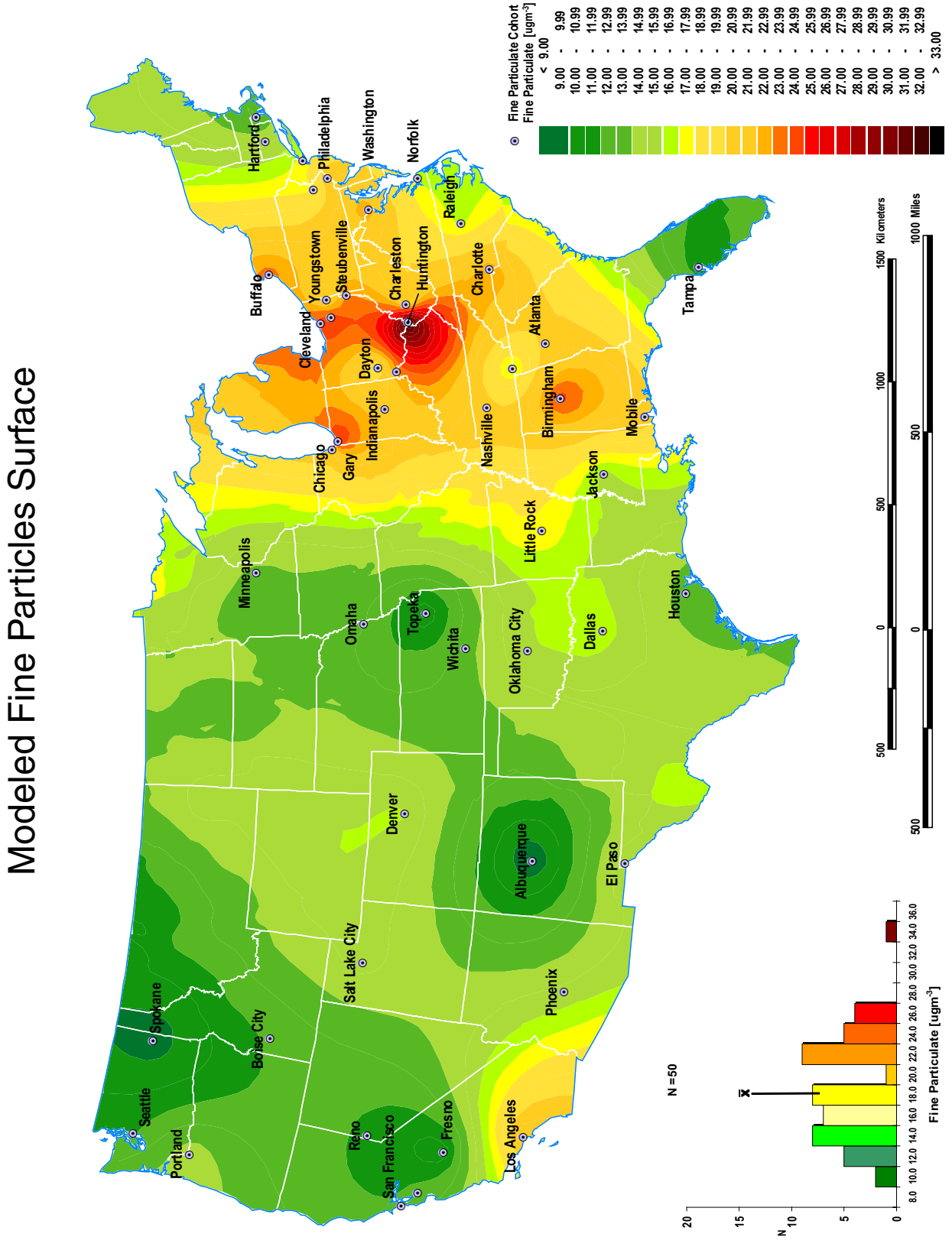

Figure 18. Spatial distribution of fine particles.

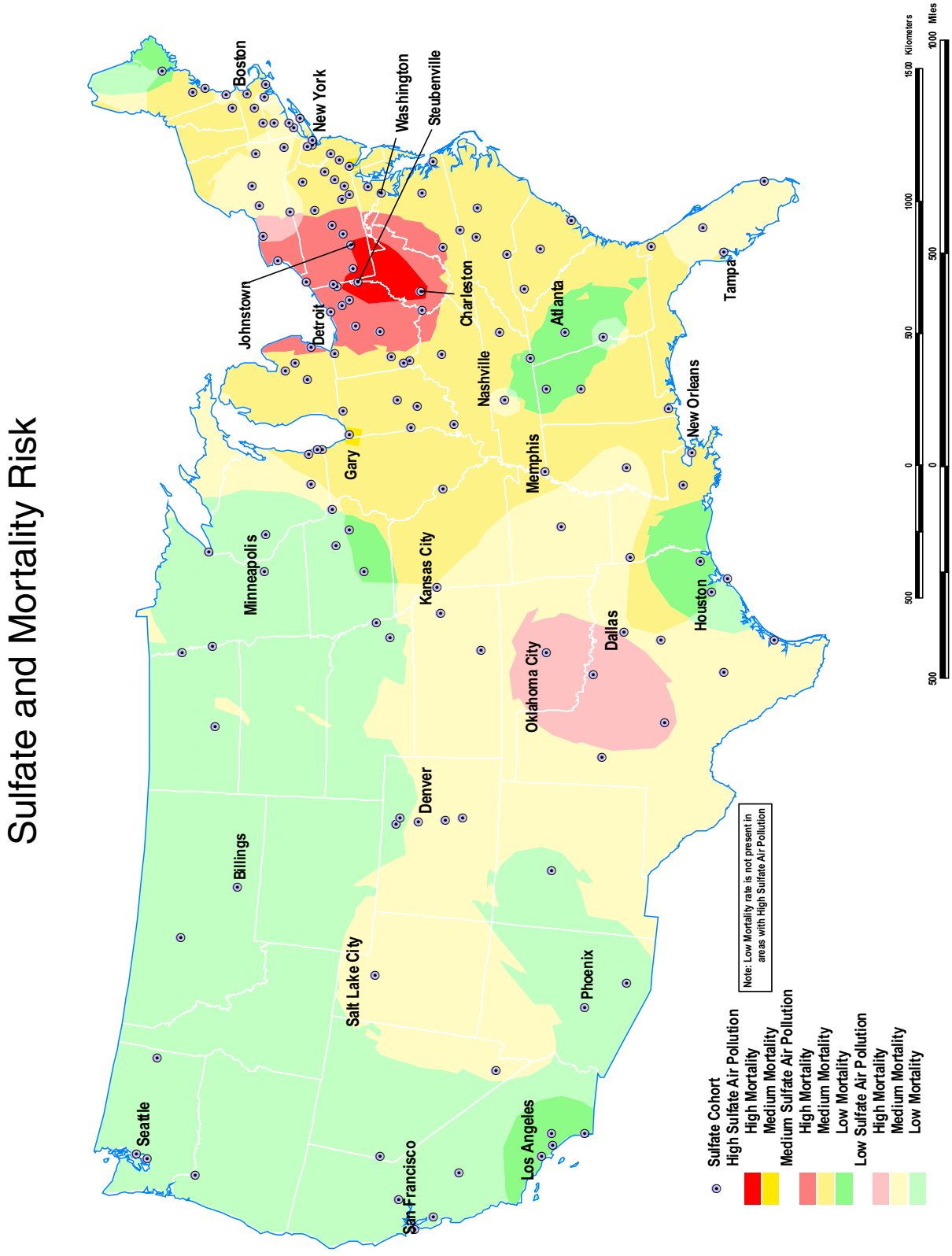

Figure 19. Spatial overlay of sulfate levels and relative risk of mortality. Interval classifications for sulfate (in $\mu g/m^3$): low 3.60–10.15; medium 10.15–16.70; high 16.70–23.25. Interval classifications for relative risks of mortality: low 0.924–1.057; medium 1.057–1.14; high 1.14–1.283.

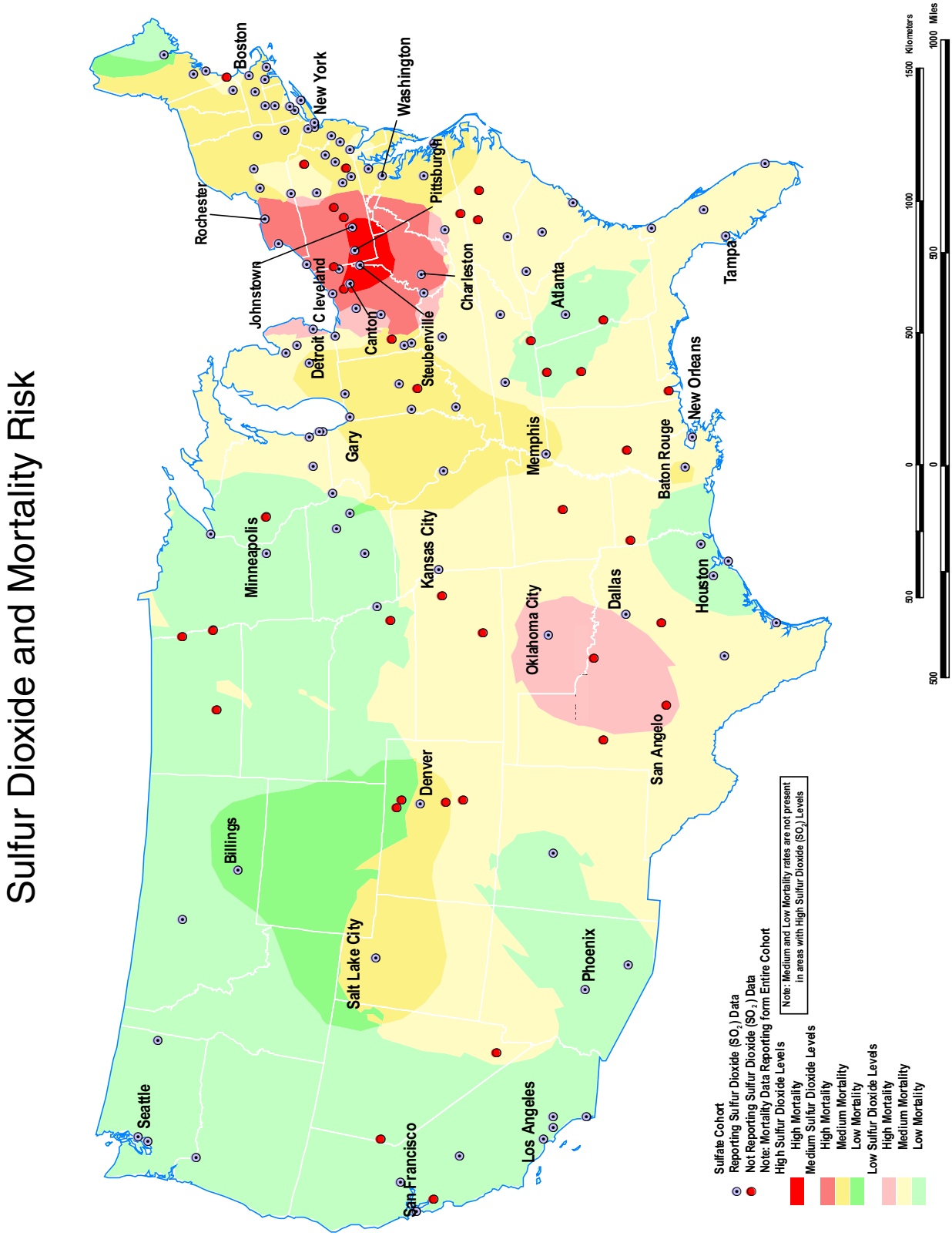

Figure 20. Spatial overlay of sulfur dioxide levels and relative risk of mortality. Interval classifications for sulfur dioxide (in ppb): low 0.05–9.78; medium 9.78–19.51; high 19.51–29.25. Interval classifications for relative risks of mortality: low 0.924–1.057; medium 1.057–1.14; high 1.14–1.283.

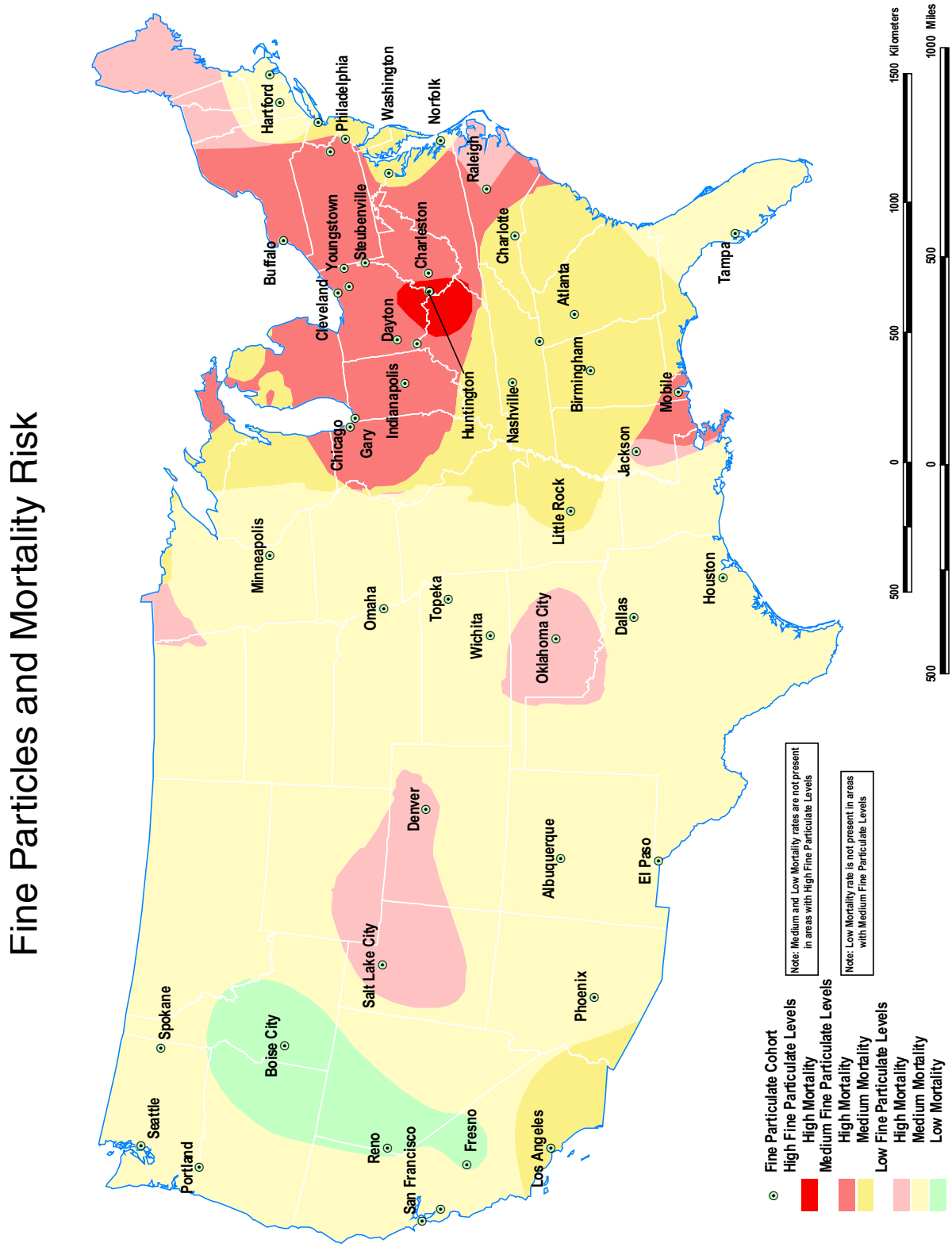

Figure 21. Spatial overlay of fine particle levels and relative risk of mortality. Interval classifications for fine particles (in $\mu g/m^3$): low 8.99–17.03; medium 17.03–25.07; high 25.07–33 Interval classifications for relative risks of mortality: low 0.502–0.711; medium 0.711–0.919; high 0.919–1.128.

these displays by first categorizing both mortality rates and pollutant levels into three equal intervals, and then selecting areas of intersection among the three categories of each variable. For the overlay of sulfate and mortality, intersections between high sulfate concentrations and high and medium mortality rates cluster mostly in the Lower Great Lakes area. (There was no intersection between high sulfate concentrations and low mortality rates.) The overlay of sulfur dioxide and mortality is similar, although there are intersections only between high mortality rates and high sulfur dioxide concentrations. (High sulfur dioxide levels do not intersect with either medium or low mortality rates.) Within the low sulfur dioxide category, a cluster of high mortality rates appears along the Texas-Oklahoma border.

The overlay of fine particles and mortality displays a similar pattern in which high levels of fine particle pollution coincide only with high mortality rates. For the medium levels of pollution, intersections exist for high and medium mortality rates, but not for low mortality rates. Only the low fine particle category intersects with the low mortality rate category. The intersection of high fine particle air pollution and high mortality rate is centered on Huntington WV. Areas along the Texas-Oklahoma border display an intersection between low fine particle air pollution and high mortality rate. Overall, these graphic results suggest a certain degree of spatial concordance between high air pollution levels and high mortality rates.

Testing for Spatial Autocorrelation

We performed exploratory tests for spatial autocorrelation in the response variables (mortality) and predictor variables (air pollution and other ecologic covariates) using global and local indicators of spatial autocorrelation statistics, including the global Moran *I* and *G* statistics (Getis and Ord 1996). We ran these tests using the S-PLUS 2000 spatial statistics package and macro programs developed by Sawada (1999) for Excel 97.

Global autocorrelation tests such as Moran *I* measure the tendency, across all metropolitan areas, for higher (or lower) values to cluster in space with other higher (or lower) values. We evaluated significance against the result expected if the data were randomly distributed in space. Positive correlations with significant *P* values suggest that high values in one metropolitan area tend to depend on values in adjacent regions (ie, higher values will cluster in space with other high values).

Global tests rely on the assumption of stationarity or structural stability over space. Nonstationarity, meaning that the relation among the attributes of interest varies spatially, is quite common. Local indicators of spatial association allow for local instabilities or nonstationarity in the data and point out areas with potential "hot spots" or clusters. The clusters indicate subregions of the study area that may have higher or lower values of attributes, such as risk of mortality, than one would expect by chance.

Results from the exploratory autocorrelation tests indicated significant spatial autocorrelation in the majority of the mortality outcomes considered in the Cox proportional-hazards model. Although they were significant, most of the Moran *I* correlation coefficients were fairly low. The results from the global Moran test indicated that mortality from all causes and from cardiopulmonary disease displays significant positive autocorrelation (Table 39) within the spatial structure of the weight matrix described in Appendix H. This suggests that high values in a given metropolitan area depend partly on other high values in adjacent metropolitan areas. In other words, high values tend to cluster together in space. Positive spatial autocorrelation usually suggests some misspecification in the original model (ie, the Cox proportional-hazards model with individual covariates), such as the omission of relevant covariates, incorrect functional form, or systematic mismeasurement of one of the variables (Odland 1988).

Spatial autocorrelation most often arises when some variable is omitted (Odland 1988), raising the possibility that part of the effect attributed to air pollution may

Table 39. Results of Global Tests for Spatial Autocorrelation in the Mortality Rate Ratios Using the Moran *I* and *G* Statistics

		Significant Hot Spots? (Moran *G*)	
Cause of Death	Moran *I*	D = 600 km	D = 440 km
All causes	0.225 ($P < 0.001$)	Yes; western lower Great Lakes to the Carolina Coast	Yes; similar but smaller pattern; not as far south
Cardiopulmonary disease	0.197 ($P < 0.001$)	Yes; lower Great Lakes east to Virginia and Maryland	Yes; similar but smaller pattern
Lung cancer	0.0307 ($P = 0.436$)	No	No

actually be due to some missing variable that is contiguous in space. We addressed this issue using spatial analytic models that incorporate the ecologic covariates assembled by the Reanalysis Team for the ACS Study.

Lung cancer mortality rates showed no significant autocorrelation within the spatial structure used for this analysis. This may be because the processes responsible for this outcome are not autocorrelated, or because the structure of the autocorrelation function differs from that corresponding to the spatial weight matrix. When the data were filtered with the local *G* statistic, we saw significant clustering in rates of all-cause mortality and cardiopulmonary disease mortality (see Table 39).

We conducted a sensitivity analysis to examine spatial dependence at lag distances ranging from 300 to 800 km. This analysis showed that 600 km was the critical distance beyond which spatial dependence decreased; consequently, we chose 600 km as the radius for the spatial filtering techniques discussed below.

Several key points arise from this analysis of spatial autocorrelation of mortality rates. Most of the relative risks of mortality display significant global autocorrelation, which needs to be taken into account in risk modeling. Local autocorrelation is significant as well, particularly in the Lower Great Lakes area. These results appear robust to alternative lag distances.

Sulfate and sulfur dioxide also show significant global autocorrelation, locally tending to cluster in the same pattern as the mortality rates. On the basis of the autocorrelation tests, individual risk factors represented in the Cox model apparently do not explain all of the observed spatial variation in mortality rate in the ACS Study data. The missing variable appears to cluster in areas of high air pollution, although it is difficult to determine whether sulfate, sulfur dioxide, or some combination of both is behind the observed spatial autocorrelation. We used the spatial regression methods discussed below to address this issue.

Two-Stage Spatial Regression Methods

The Reanalysis Team developed a two-stage regression method to take into account spatial patterns in the mortality rate data and clustering by city. In Stage 1, we regressed risk factor information at the individual level (specifically, those risks included in our Extended Model, discussed in the Alternative Risk Models section) and indicator functions for city (selecting one city as an index) using the Cox model, assuming that the data are statistically independent. For the sulfate cohort, we used Greenville SC as the index city because it had sulfate levels near the mean concentration (11 $\mu g/m^3$). We selected Raleigh NC as the index city for the fine particle cohort. After we adjusted for individual-level risk information, the logarithms of the mortality rates associated with exposure to fine particle air pollution in each city were, in fact, comparable to those in Greenville. Therefore, we note that any index city could have been selected with identical results.

In Stage 2, we regressed the logarithms of the city-specific relative mortality rates on ecologic variables that have common values for all cohort members within each city but vary among cities, including indices of air pollution and the ecologic covariates discussed previously. The statistical uncertainty in the estimated mortality rates in this step is taken into account by weighting the mortality rates by the sum of the variation in relative risks between cities and their estimation errors. (The estimation error is obtained from the Cox model.)

Our method of estimating adjusted city-specific mortality rates has the limitation that the estimate for each city was obtained as a comparison to an index city (Greenville SC). This induces additional covariation in the city-specific mortality estimates and inflates the variance of each estimate. We corrected for this variance inflation using an approach suggested by Easton and colleagues (1991). This approach also yields a variance estimate for the index city, so it can be used in the second step of the analysis, which requires a weight to be assigned to the city-specific mortality rates.

Because the covariance between the estimates of the city-specific mortality rates was almost identical for each city, the Easton adjustment is greatly simplified. The variance estimate for the index city is the average of the covariances of the city-specific estimates, and the adjusted variance for each city is the difference between the variance obtained from the Cox model and the average of the covariances.

Our random effects model assumes that the logarithms of the city-specific mortality rates follow a statistical distribution, with the expected values given by a linear regression model composed of city-level variables and a dispersion parameter that represents the true variation in the logarithms of the city-specific mortality rates, after adjusting for all risk factor information at the city level. (Note we had previously adjusted the mortality rates for individual-level information in Stage 1 of the analysis.) However, we could not observe the true values of the city-specific mortality rates and used estimates from Stage 1 that contained some inherent uncertainty. This uncertainty can be incorporated in Stage 2 by assuming that the variance of the logarithm of the city-specific mortality rate estimates is given by the sum of the true variation in rates between cities and the city-specific estimation error.

Implementation of Stage 2 of the two-stage regression approach involved the following four steps.

1. We obtained an initial estimate of the true variation in mortality between cities, using a method developed by DerSimonian and Laird (1986) that employs no city-level covariates. This method assumes knowledge only of the mean and variance structure of the data and is relatively simple to implement.
2. We used weighted linear regression methods to regress the logarithms of the city-specific mortality rates on city-level covariates, weighted by the inverse of the sum of the initial estimate of the true variance in rates between cities and the city-specific estimation variances.
3. We then used residuals from the model in step 2 to obtain an updated estimate of the true variability in rates between cities.
4. Finally, we repeated step 2 using an updated estimate of between-city variance.

We also estimated the effects on mortality of various indices of air pollution and several other covariates measured at the community level, setting the variation in rates to zero and the residual variance equal to the within-city estimation error. For these analyses, we assumed that the observations were independent and not clustered by city. (This independence assumption was also made in the Cox model.) We compared the results of our two-stage estimation approach to those of the Cox model for these analyses.

Variation in mortality rates between cities incurs a weighting scheme in which the weights are more uniform between cities compared with a scheme in which we assume no such variation (ie, the Independence Cities Model). When weights are more uniform, cities with large sample sizes will carry relatively less weight in the regression model than when we assume the observations are independent. As the estimation of error of the city-specific mortality grows smaller, the weight it is assigned in the second stage of the analysis grows larger. We would assign almost equal weights to every city if the true variation in rates was much greater than the error of the city-specific mortality rate estimates. The relative magnitude of the between-city variation in mortality rates compared with the average of the within-city estimation error determines the weighting scheme used in the analysis.

The Reanalysis Team used logarithms of the estimated city-specific mortality rates adjusted for the individual-level risk factors and their corresponding variances (which include the estimate of the true variation between cities) as input to our spatial analyses. We focused on four different two-stage regression models that afforded increasing control for spatial autocorrelation (see Tables 40 through 49).

Independent Observations Model The two-stage Independent Observations Model, like the standard Cox model, assumes that all observations are statistically independent. We obtained relative risks of mortality by fitting the Cox model with an indicator variable for each city in the first stage, and then combining the city-specific relative risks in the second stage with weights proportional to the standard errors of the relative risks in the second stage. This model provides a baseline against which either model can be compared.

Independent Cities Model The Independent Cities Model, which allows for clustering in mortality rates by city, employs a random effects approach to describe between-city variation. The random effects approach avoids the assumption of independent observations by incorporating between-city variation into the second-stage weights. However, this approach assumes that the city-specific mortality rates are statistically independent, thereby ignoring possible regional patterns in mortality that extend beyond MSA boundaries.

Regional Adjustment Model To allow for the possibility of regional effects, we conducted further analyses in which an indicator variable represented each of the seven regions shown in Figure 15. We then combined these indicator variables in the second stage to allow for residual between-city variation.

Spatial Filtering Model The final analysis summarized in Table 40 uses spatial filtering techniques to remove regional patterns in both mortality and the ecologic predictors of mortality. Variation in relative risks between cities was modeled using the two-stage random effects regression approach on the filtered mortality and ecologic covariate data. In contrast, the previous Regional Adjustment Model adjusts for spatial patterns in mortality, but not in the ecologic covariates used to predict mortality (see Appendix H for a more detailed explanation). The spatial filtering approach compares the relative risk of mortality for a given city to the risks for cities within a specified distance of that city. The distance (600 km) was selected such that the residual spatial autocorrelation was minimized. The results of applying these four two-stage regression methods to the sulfate and fine particle cohorts of the ACS Study are described below.

Table 40. Impact of Individual Ecologic Covariates on the Relative Risks of Mortality from All Causes Associated with an Increase in Sulfate Using Two-Stage Spatial Analytic Methods[a]

			Random Effects					
	Independent Observations		Independent Cities		Regional Adjustment		Spatial Filtering[b]	
Ecologic Covariate	Sulfate	Covariate	Sulfate	Covariate	Sulfate	Covariate	Sulfate	Covariate
Sulfate Alone	1.17 (1.07–1.27)		1.25 (1.13–1.37)[c]		1.19 (1.06–1.34)		1.09 (1.01–1.19)	
Demographic Factors								
Population change	1.06 (1.00–1.13)	0.86 (0.81–0.90)	1.16 (1.05–1.29)[c]	0.88 (0.80–0.96)	1.17 (1.02–1.33)	0.94 (0.84–1.05)	1.10 (1.00–1.20)	1.02 (0.86–1.22)
Whites	1.21 (1.14–1.28)	1.08 (1.04–1.13)	1.27 (1.15–1.39)[d]	1.06 (0.99–1.14)	1.20 (1.06–1.36)	1.03 (0.95–1.10)	1.10 (1.01–1.20)	1.03 (0.95–1.11)
Blacks	1.19 (1.12–1.27)	0.95 (0.90–1.00)	1.26 (1.15–1.39)[c]	0.96 (0.88–1.04)	1.19 (1.06–1.35)	0.99 (0.90–1.09)	1.10 (1.01–1.20)	0.99 (0.93–1.05)
Socioeconomic Factors								
Income	1.16 (1.10–1.23)	0.88 (0.84–0.93)	1.23 (1.12–1.35)[c]	0.91 (0.83–0.99)	1.18 (1.05–1.33)	0.90 (0.83–0.97)	1.09 (1.00–1.18)	0.91 (0.84–0.98)[e]
Poverty	1.17 (1.11–1.24)	0.95 (0.91–0.99)	1.25 (1.14–1.37)[c]	0.97 (0.89–1.05)	1.19 (1.06–1.35)	0.99 (0.91–1.08)	1.09 (1.00–1.19)	1.01 (0.94–1.09)[e]
Income disparity	1.17 (1.10–1.23)	0.87 (0.82–0.92)	1.24 (1.13–1.36)[d]	0.92 (0.83–1.01)	1.19 (1.06–1.34)	0.95 (0.86–1.06)	1.10 (1.01–1.19)	0.96 (0.88–1.06)
Unemployment	1.13 (1.07–1.20)	1.13 (1.06–1.20)	1.22 (1.11–1.34)[c]	1.11 (1.01–1.22)	1.18 (1.05–1.33)	1.06 (0.97–1.17)	1.08 (1.00–1.18)	1.08 (0.98–1.19)
Education	1.13 (1.06–1.20)	0.92 (0.86–0.98)	1.20 (1.08–1.33)[c]	0.92 (0.82–1.02)	1.16 (1.02–1.31)	0.91 (0.80–1.03)	1.07 (0.98–1.17)	0.92 (0.83–1.02)
Health Services								
Physicians	1.16 (1.10–1.22)	0.92 (0.87–0.97)	1.23 (1.12–1.36)[c]	0.92 (0.84–1.02)	1.17 (1.04–1.33)	0.93 (0.85–1.02)	1.09 (1.00–1.19)	0.94 (0.86–1.03)[e]
Hospital beds	1.15 (1.08–1.21)	1.15 (1.07–1.22)	1.24 (1.13–1.36)[c]	1.13 (1.02–1.25)	1.19 (1.05–1.35)	1.13 (1.02–1.25)	1.09 (1.00–1.19)	1.08 (0.99–1.17)
Climate								
Temperature	1.12 (1.05–1.19)	0.90 (0.86–0.94)	1.23 (1.11–1.36)[c]	0.94 (0.87–1.01)	1.20 (1.05–1.37)	0.98 (0.86–1.13)	NP	NP
Temperature variation	1.11 (1.05–1.18)	1.18 (1.11–1.25)	1.22 (1.11–1.35)[c]	1.13 (1.03–1.25)	1.18 (1.03–1.34)	1.12 (0.99–1.26)	1.08 (0.99–1.18)	1.07 (0.96–1.19)
Relative humidity	1.14 (1.05–1.24)	0.99 (0.92–1.06)	1.17 (1.01–1.35)[c]	1.02 (0.90–1.16)	1.15 (0.95–1.39)	1.06 (0.90–1.24)	1.12 (0.99–1.26)	1.10 (1.03–1.18)
Relative humidity variation	1.17 (1.09–1.26)	0.91 (0.86–0.97)	1.21 (1.07–1.37)[c]	0.92 (0.83–1.03)	1.16 (0.96–1.40)	0.99 (0.87–1.12)	1.09 (0.96–1.24)	0.96 (0.86–1.08)
Physical Environment								
Altitude	1.18 (1.10–1.27)	1.15 (1.08–1.23)	1.24 (1.09–1.40)[c]	1.07 (0.95–1.20)	1.12 (0.96–1.31)[d]	1.06 (0.94–1.21)	1.08 (0.97–1.19)	1.05 (0.95–1.17)
Water hardness	1.14 (1.07–1.21)	1.09 (1.04–1.15)	1.24 (1.11–1.39)	1.07 (0.98–1.17)	1.17 (1.01–1.36)	1.05 (0.94–1.16)	1.08 (0.98–1.19)	1.08 (0.98–1.20)
Gaseous Copollutants								
CO	1.16 (1.10–1.23)	0.96 (0.91–1.02)	1.25 (1.14–1.37)[c]	0.99 (0.90–1.09)	1.17 (1.03–1.33)	1.06 (0.97–1.16)	NP	NP
NO_2	1.16 (1.08–1.24)	0.89 (0.85–0.94)	1.27 (1.13–1.44)[c]	0.93 (0.83–1.04)	1.21 (1.02–1.44)	0.94 (0.82–1.07)	NP	NP
O_3	1.17 (1.11–1.24)	0.93 (0.87–0.99)	1.26 (1.15–1.38)[c]	0.93 (0.84–1.02)	1.16 (1.03–1.30)	1.02 (0.92–1.12)	1.08 (1.00–1.18)	0.96 (0.87–1.05)[e]
SO_2	1.05 (0.98–1.12)	1.31 (1.23–1.40)	1.13 (1.02–1.25)	1.27 (1.15–1.40)	1.10 (0.97–1.24)	1.23 (1.11–1.36)	1.05 (0.97–1.14)	1.19 (1.10–1.29)

[a] Effects were evaluated for the respective range of each covariate specified in Appendix G, which is available upon request from the Health Effects Institute. Data are RRs with 95% CIs.

[b] Used the Filtered Both Sides Model. NP = Not possible to remove spatial autocorrelation in this covariate.

[c] The residuals are spatially autocorrelated ($P < 0.05$) confidence using the Moran I statistic.

[d] The residuals are spatially autocorrelated ($P < 0.10$) confidence using the Moran I statistic.

[e] No evidence of spatial autocorrelation was found in the original data for these variables, therefore the Filtered Mortality Only Model was applied.

Spatial Analysis of the Sulfate Cohort

All-Cause Mortality Using the two-stage methods described above, we calculated the relative risk of mortality from all causes associated with exposure to sulfate (Table 40). For each of the four types of analyses, the first column of results in Table 40 represents the sulfate-exposure-associated relative risk of mortality from all causes, adjusted for each of the ecologic covariates considered; the second column represents the relative risk of all-cause mortality for the ecologic covariate, adjusted for sulfate. Because we were unable to obtain information on some ecologic covariates for certain cities (see Table 33), the relative risk of mortality associated with sulfate alone varied with the set of cities for which data were available for the covariate.

With some exceptions, the relative risk of all-cause mortality associated with exposure to sulfate in each of these subsets of cities was generally comparable to that calculated using all 151 cities. Differences occurred in the Independent Cities Model, in which RR = 1.25 (95% CI: 1.13–1.37) for all-cause mortality associated with sulfate, somewhat higher than the RR = 1.17 (95% CI: 1.01–1.35) in the 95 cities for which data on relative humidity were available. Under the Regional Adjustment Model, the relative risks of mortality associated with exposure to sulfate in these same 95 cities (RR = 1.15, 95% CI: 0.95–1.39) for which relative humidity data were available and the 110 cities (RR = 1.12, 95% CI: 0.96–1.31) for which altitude was available were somewhat lower than the relative risk based on all 151 cities (RR = 1.19, 95% CI: 1.06–1.34). Under the Filtered Both Sides Model, the relative risk of all-cause mortality associated with sulfate exposure was not significant in the subsets of cities for which data were available on education (151 cities), temperature variation (135 cities), relative humidity and relative humidity variation (95 cities), altitude (110 cities), water hardness (109 cities), and SO_2 (113 cities).

Under the Independent Observations Model, applied to all 151 cities in the sulfate cohort, the relative risk of mortality from all causes (see Table 40) was estimated to be 1.17 (95% CI: 1.07–1.27). This was similar to the estimate (RR = 1.15, 95% CI: 1.09–1.21) arrived at with the Cox proportional-hazards regression model (see Table 34). The association between exposure to sulfate and all-cause mortality remained significant after adjustment for each of the individual ecologic covariates other than population change and exposure to sulfur dioxide. Whereas population change correlated negatively with mortality when analyzed as the covariate alone, exposure to sulfur dioxide demonstrated a positive association (see Table 40).

When we allowed for clustering by city in the Independent Cities Model, we obtained higher estimates of the relative risk of all-cause mortality from exposure to sulfate than we did in the Independent Observations Model. In the Independent Cities Model, the city-specific weights used in the second stage were more uniform than in the Independent Observations Model, so that larger cities are assigned less weight. In this case, the association between sulfate and mortality remained significant even after adjustment for population change.

Although the Independent Cities Model allows for clustering within cities, it does not allow for clustering at a broader regional level. To evaluate the validity of this analysis, we used the Moran I test for global spatial autocorrelation to test for regional clustering within a radius of 600 km. Except for the analyses adjusted for water hardness and sulfur dioxide, in all cases the residuals demonstrated significant ($P < 0.05$) spatial autocorrelation, indicating the need to allow for regional clustering in the analysis.

When we adjusted for spatial clustering in city-specific mortality rates (Regional Adjustment Model) using the seven regions shown in Figure 15, we obtained relative risk estimates closer to those of the Independent Observations Model, although the confidence limits were somewhat wider. This reduction in risk accompanying regional adjustment suggests that part of the apparent sulfate effect observed with the Independent Cities Model is the result of spatial concordance between mortality and air pollution. We observed little evidence of residual spatial autocorrelation after regional adjustment, indicating that the Regional Adjustment Model removes broad regional trends in the data.

The final analysis summarized in Table 40, which used spatial filtering techniques before regression analysis was applied, removed regional trends in both mortality and each of the ecologic covariates considered. This analysis provided a more complete adjustment for regional patterns in the data without the need to specify regional boundaries as in the Regional Adjustment model. The Spatial Filtering Model resulted in relative risks of all-cause mortality associated with exposure to sulfate that were lower than those from the Regional Adjustment Model. The effect of exposure to sulfate without adjustment for any of the ecologic covariates remained significant (RR = 1.09, 95% CI: 1.01–1.19) but lower than in the Independent Cities Model; again, however, it was no longer significant after adjustment was made for exposure to sulfur dioxide.

We evaluated the stability of the sulfate–mortality association to adjustment for the effects of multiple ecologic covariates by conducting three additional multivariate regression analyses. The first analysis included all four

gaseous copollutants (CO, NO_2, O_3, and SO_2), in addition to sulfate, and was intended to examine the effect of sulfate after adjusting for all of the gaseous copollutants simultaneously. The second included population change and all of the socioeconomic factors (educational attainment, income, poverty rate, income disparity, and unemployment rate) along with sulfate. The third analysis included all ecologic covariates that individually produced a 25% change in the relative risk of mortality associated with sulfate (the covariates included were different for each analysis).

Because sulfur dioxide was the only gaseous copollutant that appeared to be associated strongly with all-cause mortality (Table 41), simultaneous adjustment for all four gaseous copollutants led to sulfate-associated relative risks of mortality somewhat similar to those obtained when we adjusted for sulfur dioxide alone. We did not see a marked impact on the association between sulfate and all-cause mortality when we adjusted simultaneously for all demographic and socioeconomic variables.

Cardiopulmonary Disease Mortality Although exposure to sulfate was significantly associated with cardiopulmonary mortality (Table 42) under the Regional Adjustment Model based on all 151 cities (RR = 1.19, 95% CI: 1.06–1.34), the sulfate-associated relative risk of cardiopulmonary disease mortality was not significant in certain subsets, notably the 95 cities in which data on relative humidity were available (RR = 1.15, 95% CI: 0.88–1.39). Under the Filtered Both Sides Model, the sulfate-associated effect also was not significant in certain subsets of cities for which ecologic data were available (notably education: RR = 1.08, 95% CI: 0.96–1.22; relative humidity: RR = 1.10, 95% CI: 0.91–1.31; relative humidity variation: RR = 1.09, 95% CI: 0.90–1.31; and altitude: RR = 1.12, 95% CI: 0.96–1.29), whereas the relative risk based on all 151 cities (RR = 1.13, 95% CI: 1.01–1.27) achieved nominal statistical significance.

The relative risk of cardiopulmonary disease mortality associated with exposure to sulfate (see Table 42) obtained with the Independent Observations Model was 1.25 (95% CI: 1.12–1.39), again similar to the value of 1.25 (95% CI: 1.16–1.36 [see Table 20, Extended Model]) achieved using the Cox regression. The sulfate-associated effect remained significant after adjustment for any one of the ecologic covariates considered, including sulfur dioxide. As was the case with mortality from all causes, the relative risk of cardiopulmonary mortality associated with exposure to sulfate tended to increase when we used the Independent Cities Model but to decrease when we applied both the Regional Adjustment Model and the Spatial Filtering Models. Although we found consistent evidence of spatial autocorrelation with the Independent Cities Model, we saw little evidence of residual spatial autocorrelation after we applied the Regional Adjustment Model. Cardiopulmonary disease mortality appeared to be associated with the unemployment rate and water hardness in addition to sulfate and sulfur dioxide, although neither unemployment nor water hardness had a marked impact on the association between sulfate and cardiopulmonary mortality. The effect of sulfate was diminished somewhat in multiple covariate models (Table 43), but remained elevated even with maximal adjustment for spatial autocorrelation.

Lung Cancer Mortality Exposure to airborne sulfate was associated with lung cancer mortality (Table 44) in both the Independent Observations Model (RR = 1.31, 95% CI: 1.05–1.65) and the Independent Cities Model (RR = 1.39, 95% CI: 1.09–1.75). None of the other ecologic covariates appeared to be associated with lung cancer, nor did they appreciably alter the association between sulfate and lung cancer mortality. The relative risk of lung cancer mortality associated with exposure to sulfate remained elevated after adjustment for multiple covariates (Table 45). Lung cancer exhibited a high degree of spatial homogeneity and there was no evidence of spatial autocorrelation in the Independent Cities Model; thus no attempt was made to remove it using either the Regional Adjustment Model or Spatial Filtering Model.

Spatial Analysis of the Fine Particle Cohort

All-Cause Mortality Exposure to fine particles was associated with all-cause mortality (Table 46) under the Independent Observations Model (RR = 1.18, 95% CI: 1.03–1.35). The relative risk increased (RR = 1.29, 95% CI: 1.12–1.48) under the Independent Cities Model, but dropped (RR = 1.16, 95% CI: 0.99–1.37) after we applied the Regional Adjustment Model. We were unable to apply the Spatial Filtering Model, largely because of the limited number of cities (50) in the fine particle cohort.

Sulfur dioxide pollution appeared to be strongly associated with all-cause mortality in the fine particle cohort, as it was in the sulfate cohort. Water hardness also showed an association with all-cause mortality in the fine particle cohort, but it had little effect in the sulfate cohort. The relative risk of all-cause mortality associated with exposure to fine particles remained elevated, if not significant, in the Independent Cities and Regional Adjustment Models. The relative risk of all-cause mortality associated with exposure to fine particles was not altered greatly after adjusting for all demographic and socioeconomic covariates, although the relative risk was reduced markedly in

Table 41. Impact of Multiple Ecologic Covariates on the Relative Risks of Mortality from All Causes Associated with an Increase in Sulfate Using Two-Stage Spatial Analytic Methods[a]

Spatial Analytic Model	Number of Cities	Relative Risk Calculated for Sulfate Alone	Ecologic Covariates Incorporated into Adjusted Analyses	Relative Risk After Adjusting for Ecologic Covariates[b]	
				Sulfate	SO_2
Demographic and Socioeconomic Factors					
Independent Observations	139	1.16 (1.10–1.23)	Population change, income, poverty, income disparity, unemployment, education	1.10 (1.02–1.18)	NC
Independent Cities	139	1.24 (1.13–1.37)		1.17 (1.05–1.31)	NC
Regional Adjustment	139	1.18 (1.04–1.34)		1.21 (1.06–1.38)	NC
Spatial Filtering[c]	139	1.10 (1.00–1.20)		1.11 (1.01–1.21)[d]	NC
Gaseous Copollutants					
Independent Observations	58	1.11 (1.04–1.19)	CO, NO_2, O_3, SO_2	1.06 (0.98–1.14)	1.41 (1.31–1.52)
Independent Cities	58	1.25 (1.10–1.43)		1.05 (0.93–1.18)	1.39 (1.24–1.55)
Regional Adjustment	58	1.25 (1.03–1.51)		1.06 (0.90–1.26)	1.28 (1.12–1.46)
Spatial Filtering[c]	102	1.09 (0.99–1.19)	O_3, SO_2	1.05 (0.96–1.14)	1.19 (1.09–1.29)
25%[e]					
Independent Observations	44	1.12 (1.03–1.21)	Population change, whites, temperature variation, relative humidity variation, altitude, SO_2	1.18 (1.07–1.30)	1.22 (1.09–1.37)
Independent Cities	103	1.28 (1.16–1.42)	Population change, SO_2	1.10 (0.99–1.22)	1.23 (1.11–1.37)
Regional Adjustment	113	1.21 (1.07–1.37)	SO_2	1.10 (0.97–1.24)	1.23 (1.11–1.36)
Spatial Filtering[c]	50	1.05 (0.91–1.22)	Education, relative humidity, altitude, SO_2	1.09 (0.94–1.26)	1.07 (0.95–1.22)

[a] Effects were evaluated for the respective range of each covariate specified in Appendix G, which is available upon request from the Health Effects Institute. Data are RRs with 95% CIs.

[b] NC = Relative risks associated with SO_2 were not calculated in some multivariate models.

[c] Used the Filtered Both Sides Model.

[d] The residuals are spatially autocorrelated ($P < 0.10$) using the Moran I statistic.

[e] Incorporated all ecologic covariates that, when analyzed individually in a bivariate model, were found to produce a change of 25% or more in the excess relative risk associated with sulfate.

Table 42. Impact of Individual Ecologic Covariates on the Relative Risks of Mortality from Cardiopulmonary Disease Associated with an Increase in Sulfate Using Two-Stage Spatial Analytic Methods[a]

			Random Effects					
	Independent Observations		Independent Cities		Regional Adjustment		Spatial Filtering[b]	
Ecologic Covariate	Sulfate	Covariate	Sulfate	Covariate	Sulfate	Covariate	Sulfate	Covariate
Sulfate Alone	1.25 (1.12–1.39)		1.29 (1.15–1.46)[c]		1.19 (1.06–1.34)		1.13 (1.01–1.27)	
Demographic Factors								
Population change	1.11 (1.01–1.21)	0.82 (0.76–0.88)	1.17 (1.03–1.33)[c]	0.84 (0.76–0.94)	1.16 (0.98–1.37)	0.91 (0.78–1.05)	1.12 (1.00–1.25)	1.08 (0.85–1.37)
Whites	1.30 (1.20–1.42)	1.10 (1.04–1.17)	1.32 (1.17–1.50)[c]	1.07 (0.98–1.18)	1.24 (1.05–1.46)	1.03 (0.93–1.13)	1.14 (1.02–1.28)	1.04 (0.94–1.16)
Blacks	1.28 (1.18–1.40)	0.94 (0.87–1.01)	1.31 (1.15–1.49)[c]	0.96 (0.86–1.07)	1.22 (1.04–1.44)	1.01 (0.89–1.14)	1.14 (1.01–1.28)	0.99 (0.91–1.07)
Socioeconomic Factors								
Income	1.25 (1.15–1.35)	0.86 (0.80–0.92)	1.27 (1.13–1.43)[c]	0.87 (0.78–0.97)	1.21 (1.04–1.42)	0.86 (0.77–0.96)	1.13 (1.01–1.26)	0.87 (0.78–0.96)[e]
Poverty	1.25 (1.16–1.36)	0.97 (0.91–1.04)	1.29 (1.14–1.46)[c]	1.00 (0.91–1.09)	1.22 (1.03–1.43)	1.04 (0.93–1.17)	1.12 (1.00–1.26)	1.06 (0.96–1.16)[e]
Income disparity	1.25 (1.16–1.35)	0.89 (0.82–0.96)	1.29 (1.14–1.46)[c]	0.94 (0.83–1.07)	1.22 (1.04–1.44)	1.00 (0.87–1.16)	1.16 (1.03–1.30)	1.02 (0.90–1.16)
Unemployment	1.19 (1.09–1.29)	1.24 (1.14–1.36)	1.24 (1.10–1.40)[c]	1.21 (1.07–1.36)	1.20 (1.02–1.40)	1.15 (1.01–1.30)	1.10 (0.99–1.23)	1.19 (1.05–1.36)
Education	1.20 (1.10–1.31)	0.90 (0.82–0.99)	1.22 (1.06–1.39)[c]	1.50 (1.14–1.98)	1.17 (0.99–1.38)	0.85 (0.72–1.00)	1.08 (0.96–1.22)	0.86 (0.74–0.99)
Health Services								
Physicians	1.23 (1.13–1.33)	0.86 (0.79–0.93)	1.25 (1.11–1.41)[c]	0.85 (0.75–0.96)	1.17 (1.00–1.37)	0.85 (0.76–0.96)	1.10 (0.98–1.23)	0.85 (0.76–0.96)[e]
Hospital beds	1.22 (1.13–1.32)	1.16 (1.06–1.28)	1.27 (1.13–1.44)[c]	1.15 (1.01–1.30)	1.20 (1.02–1.41)	1.14 (0.99–1.30)	1.11 (0.99–1.25)	1.08 (0.96–1.20)
Climate								
Temperature	1.18 (1.08–1.28)	0.88 (0.83–0.94)	1.25 (1.09–1.42)[c]	0.93 (0.84–1.03)	1.21 (1.01–1.44)	0.97 (0.81–1.17)	NP	NP
Temperature variation	1.16 (1.07–1.26)	1.25 (1.15–1.35)	1.23 (1.08–1.40)[c]	1.19 (1.05–1.36)	1.17 (0.98–1.39)[d]	1.19 (1.01–1.39)	1.10 (0.97–1.24)	1.11 (0.96–1.28)
Relative humidity	1.17 (1.04–1.33)	1.03 (0.92–1.15)	1.17 (0.96–1.43)[c]	1.05 (0.88–1.26)	1.15 (0.88–1.39)	1.11 (0.89–1.38)	1.10 (0.91–1.31)	1.09 (0.98–1.20)
Relative humidity variation	1.26 (1.14–1.40)	0.86 (0.79–0.95)	1.25 (1.06–1.48)[c]	0.89 (0.76–1.03)	1.17 (0.91–1.52)	0.97 (0.81–1.16)	1.09 (0.90–1.31)	0.93 (0.79–1.09)
Physical Environment								
Altitude	1.27 (1.15–1.40)	1.15 (1.04–1.26)	1.28 (1.09–1.52)[c]	1.04 (0.89–1.22)	1.23 (0.99–1.53)	1.01 (0.84–1.21)	1.12 (0.96–1.29)	1.02 (0.88–1.19)
Water hardness	1.20 (1.10–1.31)	1.13 (1.05–1.21)	1.26 (1.10–1.44)[d]	1.11 (1.00–1.24)	1.14 (0.96–1.37)	1.08 (0.95–1.22)	1.11 (0.98–1.26)	1.11 (0.97–1.26)
Gaseous Copollutants								
CO	1.28 (1.18–1.39)	0.93 (0.86–1.01)	1.32 (1.18–1.48)[c]	0.94 (0.84–1.05)	1.24 (1.06–1.45)	0.98 (0.88–1.10)	NP	NP
NO_2	1.29 (1.17–1.42)	0.89 (0.82–0.96)	1.35 (1.18–1.55)[c]	0.91 (0.81–1.02)	1.32 (1.07–1.62)	0.92 (0.79–1.07)	NP	NP
O_3	1.27 (1.17–1.38)	0.96 (0.88–1.06)	1.32 (1.18–1.48)[c]	0.96 (0.85–1.09)	1.18 (1.01–1.37)	1.08 (0.96–1.23)	1.13 (1.01–1.26)	0.99 (0.88–1.12)[e]
SO_2	1.13 (1.03–1.24)	1.35 (1.23–1.47)	1.18 (1.04–1.34)	1.30 (1.16–1.47)	1.12 (0.96–1.32)	1.27 (1.12–1.44)	1.10 (0.99–1.22)	1.23 (1.11–1.37)

[a] Effects were evaluated for the respective range of each covariate specified in Appendix G, which is available upon request from the Health Effects Institute. Data are RRs with 95% CIs.

[b] Used the Filtered Both Sides Model. NP = Not possible to remove spatial autocorrelation in this covariate.

[c] The residuals are spatially autocorrelated ($P < 0.05$) using the Moran I statistic.

[d] The residuals are spatially autocorrelated ($P < 0.10$) using the Moran I statistic.

[e] No evidence of spatial autocorrelation was found in the original data for these variables, therefore the Filtered Mortality Only Model was applied.

Table 43. Impact of Multiple Ecologic Covariates on the Relative Risks of Mortality from Cardiopulmonary Disease Associated with an Increase in Sulfate Using Two-Stage Spatial Analytic Methods[a]

Spatial Analytic Model	Number of Cities	Relative Risk Calculated for Sulfate Alone	Ecologic Covariates Incorporated into Adjusted Analyses	Relative Risk After Adjusting for Ecologic Covariates[b]	
				Sulfate	SO_2
Demographic and Socioeconomic Factors					
Independent Observations	139	1.24 (1.15–1.34)	Population change, income, poverty, income disparity, unemployment, education	1.15 (1.04–1.28)	NC
Independent Cities	139	1.28 (1.13–1.45)		1.18 (1.02–1.37)[c]	NC
Regional Adjustment	139	1.19 (1.01–1.40)		1.21 (1.01–1.44)	NC
Spatial Filtering[d]	139	1.12 (1.00–1.26)		1.12 (0.99–1.27)	NC
Gaseous Copollutants					
Independent Observations	58	1.23 (1.11–1.36)	CO, NO_2, O_3, SO_2	1.11 (0.99–1.24)	1.43 (1.29–1.59)
Independent Cities	58	1.31 (1.13–1.52)		1.11 (0.97–1.27)	1.42 (1.25–1.61)
Regional Adjustment	58	1.36 (1.08–1.72)		1.15 (0.93–1.42)	1.30 (1.11–1.52)
Spatial Filtering[d]	102	1.14 (1.02–1.28)	O_3, SO_2	1.10 (0.99–1.23)	1.33 (1.17–1.51)
25%[e]					
Independent Observations	45	1.21 (1.07–1.36)	Population change, unemployment, temperature variation, relative humidity variation, altitude, SO_2	1.02 (0.84–1.25)	1.37 (1.16–1.62)
Independent Cities	103	1.34 (1.18–1.51)	Population change, education, SO_2	1.07 (0.93–1.24)	1.29 (1.14–1.46)
Regional Adjustment	113	1.25 (1.07–1.47)	SO_2	1.12 (0.96–1.32)	1.18 (1.08–1.28)
Spatial Filtering[d]	72	1.10 (0.92–1.31)	Education, relative humidity, relative humidity variation, SO_2	1.20 (1.01–1.43)	1.19 (1.02–1.38)

[a] Effects were evaluated for the respective range of each covariate specified in Appendix G, which is available upon request from the Health Effects Institute. Data are RRs with 95% CIs.

[b] NC = Relative risks associated with SO_2 were not calculated in some multivariate models.

[c] The residuals are spatially autocorrelated ($P < 0.05$) using the Moran I statistic.

[d] Used the Filtered Both Sides Model.

[e] Incorporated all ecologic covariates that, when analyzed individually in a bivariate model, were found to produce a change of 25% or more in the excess relative risk associated with sulfate.

Table 44. Impact of Individual Ecologic Covariates on the Relative Risks of Mortality from Lung Cancer Associated with an Increase in Sulfate Using Two-Stage Spatial Analytic Methods[a]

	Independent Observations		Random Effects Independent Cities[b]	
Ecologic Covariate	Sulfate	Covariate	Sulfate	Covariate
Sulfate Alone	1.31 (1.05–1.65)		1.39 (1.09–1.75)	
Demographic Factors				
Population change	1.27 (1.03–1.57)	0.97 (0.82–1.15)	1.36 (1.04–1.77)	0.99 (0.80–1.23)
Whites	1.33 (1.10–1.62)	1.04 (0.91–1.20)	1.39 (1.09–1.78)	1.02 (0.85–1.22)
Blacks	1.35 (1.11–1.65)	0.93 (0.79–1.10)	1.40 (1.10–1.80)	0.96 (0.78–1.18)
Socioeconomic Factors				
Income	1.31 (1.08–1.58)	0.98 (0.84–1.15)	1.39 (1.09–1.76)	1.01 (0.82–1.24)
Poverty	1.32 (1.09–1.59)	1.03 (0.89–1.20)	1.39 (1.09–1.76)	0.99 (0.81–1.20)
Income disparity	1.30 (1.08–1.57)	0.87 (0.72–1.05)	1.38 (1.09–1.75)	0.90 (0.71–1.15)
Unemployment	1.27 (1.05–1.55)	1.13 (0.93–1.38)	1.34 (1.06–1.71)	1.15 (0.90–1.46)
Education	1.30 (1.06–1.61)	0.98 (0.79–1.23)	1.38 (1.06–1.80)	1.00 (0.75–1.31)
Health Services				
Physicians	1.27 (1.05–1.54)	0.90 (0.74–1.10)	1.35 (1.06–1.71)	0.92 (0.72–1.19)
Hospital beds	1.30 (1.07–1.57)	0.89 (0.71–1.13)	1.37 (1.08–1.75)	0.87 (0.66–1.15)
Climate				
Temperature	1.34 (1.10–1.64)	0.98 (0.84–1.14)	1.43 (1.12–1.84)	0.99 (0.82–1.21)
Temperature variation	1.37 (1.12–1.68)	0.95 (0.78–1.15)	1.46 (1.14–1.88)	0.92 (0.72–1.18)
Relative humidity	1.33 (1.00–1.78)	1.22 (0.93–1.58)	1.34 (0.96–1.87)	1.21 (0.90–1.63)
Relative humidity variation	1.45 (1.13–1.87)	1.11 (0.89–1.39)	1.47 (1.10–1.97)	1.08 (0.84–1.40)
Physical Environment				
Altitude	1.11 (0.89–1.40)	0.76 (0.58–0.99)	1.13 (0.86–1.50)	0.75 (0.56–1.01)
Water hardness	1.26 (1.03–1.53)	0.94 (0.79–1.13)	1.41 (1.08–1.86)	0.94 (0.75–1.18)
Gaseous Copollutants				
CO	1.26 (1.03–1.53)	0.82 (0.67–0.99)	1.29 (1.03–1.61)	0.83 (0.66–1.03)
NO_2	1.31 (1.05–1.65)	0.87 (0.72–1.05)	1.36 (1.04–1.76)	0.88 (0.71–1.10)
O_3	1.30 (1.07–1.59)	0.71 (0.53–0.96)	1.33 (1.07–1.65)	0.72 (0.57–0.91)
SO_2	1.37 (1.08–1.73)	0.94 (0.75–1.18)	1.39 (1.08–1.81)	0.94 (0.73–1.20)

[a] Effects were evaluated for the respective range of each covariate specified in Appendix G, which is available upon request from the Health Effects Institute. Data are RRs with 95% CIs.

[b] Neither the raw lung cancer relative risks nor the residuals of the Independent Cities Model incorporating spatially autocorrelated sulfate and covariate values were found to be spatially autocorrelated. Therefore, it was not necessary to analyze Regional Adjustment or Spatial Filtering Models.

Table 45. Impact of Multiple Ecologic Covariates on the Relative Risks of Mortality from Lung Cancer Associated with an Increase in Sulfate Using Two-Stage Spatial Analytic Methods[a]

Spatial Analytic Model	Number of Cities	Relative Risk Calculated for Sulfate Alone	Ecologic Covariates Incorporated into Adjusted Analyses	Relative Risk After Adjusting for Ecologic Covariates[b]	
				Sulfate	SO_2
Demographic and Socioeconomic Factors					
Independent Observations	139	1.29 (1.07–1.56)	Population change, income, poverty, income disparity, unemployment, education	1.14 (0.89–1.45)	NC
Independent Cities	139	1.36 (1.07–1.74)		1.23 (0.90–1.68)	NC
Gaseous Copollutants					
Independent Observations	58	1.42 (1.11–1.80)	CO, NO_2, O_3, SO_2	1.61 (1.21–2.15)	0.87 (0.65–1.17)
Independent Cities	58	1.48 (1.12–1.96)		1.63 (1.19–2.23)	0.90 (0.67–1.21)
25%[c]					
Independent Observations	68	1.61 (1.22–2.14)	Relative humidity, altitude	1.39 (0.98–1.99)	NC
Independent Cities	68	1.62 (1.21–2.16)		1.39 (0.97–2.01)	NC

[a] Effects were evaluated for the respective range of each covariate specified in Appendix G, which is available upon request from the Health Effects Institute. Data are RRs with 95% CIs.

[b] NC = Relative risks associated with SO_2 were not calculated in some multivariate models.

[c] Incorporated all ecologic covariates that, when analyzed individually in a bivariate model, were found to produce a change of 25% or more in the excess relative risk associated with sulfate.

multiple covariate models that included sulfur dioxide as a copollutant (Table 47). (In contrast, the effect of exposure to sulfur dioxide persisted even in the three multiple covariate analyses considered.)

Cardiopulmonary Disease Mortality Fine particle air pollution alone was associated with cardiopulmonary mortality (Table 48) under all three models considered, with relative risks of 1.30, 1.38, and 1.24 under the Independent Observations, Independent Cities, and Regional Adjustment Models, respectively. In the fine particle cohort, as in the sulfate cohort, unemployment appeared to be associated with cardiopulmonary mortality, although adjustment for unemployment rate in the sulfate cohort did not have a marked impact on the relative risk. Sulfur dioxide was strongly associated with cardiopulmonary disease mortality, although the fine particle effect on cardiopulmonary disease mortality was not eliminated by adjustment for exposure to sulfur dioxide. After we applied the Regional Adjustment Model, there was no evidence of residual spatial autocorrelation in cardiopulmonary mortality as there had been for all-cause mortality. Multivariate adjustment (Table 49) reduced, but did not eliminate, the fine particle–cardiopulmonary association.

Lung Cancer Mortality Because we detected no association between exposure to fine particles and lung cancer mortality using Cox regression, we conducted no further spatial analyses.

Simultaneous Autoregressive Models

In the preceding section, we used two approaches to adjust for broad spatial patterns in the ACS data. In the first, we used the Regional Adjustment Model to remove spatial variation in mortality rates, adjusting the city-specific values for broad regional patterns. In the second, we used spatial filtering techniques to remove spatial patterns in both the mortality data and the city-level variables before we linked them together. With this analytic approach, after broad regional patterns have been removed, attributes of both predictor and response variables are compared using random effects regression methods.

We further examined the robustness of our results to the method of controlling for spatial autocorrelation by using a third modeling approach, namely the Simultaneous Autoregressive Model described in Appendix H. In this approach the logarithms of the city-specific mortality rates are the response variables and are assumed to be normally distributed. City-level covariates are included as predictors of mortality; however, the error structure incorporates the spatial autocorrelation between mortality rates after accounting for city-level predictors of mortality.

Table 46. Impact of Individual Ecologic Covariates on the Relative Risks of Mortality from All Causes Associated with an Increase in Fine Particles Using Two-Stage Spatial Analytic Methods[a]

			Random Effects			
	Independent Observations		Independent Cities		Regional Adjustment	
Ecologic Covariate	Fine Particles	Covariate	Fine Particles	Covariate	Fine Particles	Covariate
Fine Particles Alone	1.18 (1.03–1.35)		1.29 (1.12–1.48)[b]		1.16 (0.99–1.37)	
Demographic Factors						
Population change	1.07 (0.98–1.17)	0.86 (0.81–0.92)	1.19 (1.01–1.39)	0.88 (0.77–0.99)	1.18 (0.97–1.42)[c]	1.00 (0.85–1.16)
Whites	1.28 (1.17–1.40)	1.15 (1.08–1.23)	1.33 (1.16–1.53)[c]	1.12 (0.99–1.26)	1.19 (1.01–1.41)[c]	1.07 (0.96–1.20)
Blacks	1.27 (1.16–1.38)	0.88 (0.82–0.96)	1.34 (1.16–1.56)[c]	0.90 (0.78–1.03)	1.19 (1.00–1.41)	0.94 (0.81–1.10)
Socioeconomic Factors						
Income	1.18 (1.10–1.27)	0.92 (0.86–0.98)	1.28 (1.11–1.47)[b]	0.96 (0.84–1.10)	1.15 (0.99–1.33)[b]	0.86 (0.77–0.96)
Poverty	1.23 (1.14–1.32)	0.87 (0.80–0.94)	1.32 (1.15–1.51)[c]	0.89 (0.78–1.02)	1.15 (0.97–1.37)[c]	1.03 (0.89–1.21)
Income disparity	1.23 (1.14–1.32)	0.83 (0.78–0.88)	1.30 (1.15–1.48)	0.84 (0.75–0.93)	1.17 (0.99–1.38)	0.94 (0.83–1.07)
Unemployment	1.16 (1.08–1.25)	1.07 (1.00–1.14)	1.28 (1.11–1.48)[b]	1.02 (0.90–1.15)	1.16 (0.98–1.37)[c]	1.02 (0.91–1.13)
Education	1.18 (1.09–1.28)	0.99 (0.93–1.06)	1.32 (1.12–1.55)[b]	1.03 (0.90–1.18)	1.14 (0.97–1.35)[c]	0.93 (0.81–1.07)
Health Services						
Physicians	1.19 (1.11–1.28)	0.94 (0.87–1.02)	1.30 (1.12–1.51)[b]	1.00 (0.85–1.18)	1.18 (0.98–1.41)[c]	0.99 (0.86–1.14)
Hospital beds	1.19 (1.10–1.28)	1.02 (0.93–1.13)	1.30 (1.13–1.50)[b]	1.02 (0.87–1.21)	1.18 (0.98–1.42)[c]	1.03 (0.88–1.21)
Climate						
Temperature	1.12 (1.03–1.21)	0.86 (0.81–0.91)	1.22 (1.08–1.39)	0.85 (0.77–0.94)	1.14 (0.98–1.33)	1.02 (0.86–1.20)
Temperature variation	1.08 (0.99–1.17)	1.16 (1.09–1.24)	1.19 (1.03–1.36)[c]	1.15 (1.02–1.29)	1.11 (0.94–1.31)	1.06 (0.94–1.20)
Relative humidity	1.18 (1.08–1.30)	1.00 (0.93–1.07)	1.23 (1.05–1.44)[b]	1.04 (0.91–1.19)	1.10 (0.91–1.34)	0.86 (0.69–1.07)
Relative humidity variation	1.21 (1.10–1.33)	0.93 (0.86–1.00)	1.26 (1.08–1.47)[c]	0.96 (0.83–1.10)	1.14 (0.94–1.39)	0.99 (0.81–1.20)
Physical Environment						
Altitude	1.14 (1.05–1.24)	1.10 (1.03–1.17)	1.21 (1.02–1.43)	1.03 (0.90–1.17)	1.09 (0.91–1.31)	1.02 (0.89–1.16)
Water hardness	1.16 (1.08–1.25)	1.14 (1.07–1.22)	1.28 (1.11–1.49)[b]	1.13 (1.00–1.29)	1.17 (0.98–1.40)[c]	1.08 (0.94–1.23)
Gaseous Copollutants						
CO	1.18 (1.10–1.27)	0.92 (0.87–0.98)	1.28 (1.10–1.48)[c]	0.95 (0.83–1.09)	1.17 (0.98–1.39)[b]	0.97 (0.85–1.10)
NO_2	1.21 (1.11–1.33)	0.91 (0.85–0.97)	1.33 (1.12–1.58)	0.94 (0.81–1.09)	1.18 (1.00–1.40)	0.93 (0.80–1.08)
O_3	1.19 (1.10–1.28)	0.89 (0.83–0.96)	1.27 (1.11–1.46)[b]	0.92 (0.80–1.06)	1.11 (0.96–1.28)	1.12 (1.00–1.26)
SO_2	1.03 (0.95–1.13)	1.46 (1.31–1.62)	1.14 (0.98–1.32)	1.40 (1.17–1.67)	1.11 (0.93–1.33)	1.24 (1.04–1.48)

[a] Effects were evaluated for the respective range of each covariate specified in Appendix G, which is available upon request from the Health Effects Institute. Data are RRs with 95% CIs.

[b] The residuals are spatially autocorrelated ($P < 0.05$) using the Moran I statistic.

[c] The residuals are spatially autocorrelated ($P < 0.10$) using the Moran I statistic.

The correlation structure is based on the nearest-neighbor concept, which assumes that a city is more influenced by its nearest neighbor than by any other city, no matter how far away the nearest neighbor is. A city's neighbors are defined in the following manner. First, each of the 151 cities is assigned a Thiessen polygon, a geographic area within which all points within the polygon are closer to the city enclosed than to any other city. Then the neighbors of any city are determined as those in all the Thiessen polygons touching the polygon of that city. Each city may have a different number of neighbors, and the nearest neighbor will be a different distance away for each city. We derived a correlation structure in which a city's residual response correlates only with the residual responses of its neighbors. Cities that are not neighbors are not assumed to be correlated. We assumed a common correlation parameter for the entire dataset and estimated it simultaneously with the regression parameters using maximum likelihood techniques in S-PLUS. We also weighted the analysis by the inverse of the sum of the estimate of the variation in mortality rates between cities and the estimation error for a given city, thus incorporating the concept of a random effects model in the analysis.

We also considered a modified nearest-neighbor modeling approach in which we assumed mortality rates

Table 47. Impact of Multiple Ecologic Covariates on the Relative Risks of Mortality from All Causes Associated with an Increase in Fine Particles Using Two-Stage Spatial Analytic Methods[a]

Spatial Analytic Model	Number of Cities	Relative Risk Calculated for Fine Particles Alone	Ecologic Covariates Incorporated into Adjusted Analyses	Relative Risk After Adjusting for Ecologic Covariates[b]	
				Fine Particles	SO_2
Demographic and Socioeconomic Factors					
Independent Observations	48	1.19 (1.11–1.28)	Population change, income, poverty, income disparity, unemployment, education	1.15 (1.03–1.27)	NC
Independent Cities	48	1.30 (1.13–1.50)		1.23 (1.02–1.48)	NC
Regional Adjustment	48	1.18 (0.98–1.41)		1.15 (0.96–1.39)	NC
Gaseous Copollutants					
Independent Observations	28	1.15 (1.05–1.26)	CO, NO_2, O_3, SO_2	1.06 (0.95–1.18)	1.48 (1.33–1.65)
Independent Cities	28	1.31 (1.10–1.56)		1.11 (0.95–1.29)	1.44 (1.23–1.69)
Regional Adjustment	28	1.18 (0.99–1.40)		1.09 (0.92–1.29)	1.19 (0.99–1.44)
25%[c]					
Independent Observations	22	1.09 (0.99–1.21)	Population change, whites, temperature variation, altitude, NO_2, SO_2	1.12 (0.96–1.31)	1.16 (0.97–1.39)
Independent Cities	32	1.32 (1.12–1.54)	Population change, temperature variation, SO_2	1.06 (0.89–1.26)	1.28 (1.04–1.57)
Regional Adjustment	27	1.21 (0.98–1.50)	Relative humidity, SO_2	1.05 (0.85–1.30)	1.25 (0.97–1.61)

[a] Effects were evaluated for the respective range of each covariate specified in Appendix G, which is available upon request from the Health Effects Institute. Data are RRs with 95% CIs.

[b] NC = Relative risks associated with SO_2 were not calculated in some multivariate models.

[c] Incorporated all ecologic covariates that, when analyzed individually in a bivariate model, were found to produce a change of 25% or more in the excess relative risk associated with fine particles.

among cities were correlated when the cities were nearest neighbors or were within the average distance between cities (111 km for the cities with sulfate data and 123 km for the cities with sulfur dioxide data). We report the results obtained using the nearest-neighbor approach only, because the results using the modified nearest-neighbor approach were almost identical. The data used in the latter to generate the correlation matrix incorporated more cities in the Northeast and Ohio Valley regions; however, the inclusion of these additional cities did not influence the estimate of the common correlation parameter and thus had little impact on our estimates of the effects on mortality of exposure to air pollution.

Using the nearest-neighbor approach, our estimated relative risk of all-cause mortality associated with exposure to sulfate (RR = 1.20, 95% CI: 1.06–1.36) was similar to that obtained from the Independent Cities Model (RR = 1.25, 95% CI: 1.13–1.37), which assumes geographic independence, or after applying the Regional Adjustment Model (RR = 1.19, 95% CI: 1.06–1.34). However, we obtained a somewhat lower relative risk of all-cause mortality (RR = 1.09, 95% CI: 1.01–1.19) when we subjected both mortality rates and sulfate levels to spatial filtering techniques. The relative risk of mortality from exposure to sulfur dioxide under the Simultaneous Autoregressive Model was 1.35 (95% CI: 1.16–1.57), a value similar to those obtained by the other methods of analysis considered (Independent Cities Model: RR = 1.33, 95% CI: 1.22–1.45; Regional Adjustment Model: RR = 1.26, 95% CI: 1.15–1.39; Filtered Both Sides Model: RR = 1.27, 95% CI: 1.15–1.40). Note that the range in excess risk [(1 – RR) × 100] was similar for sulfate (24% – 12% = 12%) and sulfur dioxide (35% – 25% = 10%), which suggests that each pollutant was equally sensitive, in absolute terms, to the statistical approach used.

When we modeled sulfate jointly with sulfur dioxide using the Simultaneous Autoregressive Model, the relative risk of mortality from exposure to sulfate was 1.08 (95%

Table 48. Impact of Individual Ecologic Covariates on the Relative Risks of Mortality from Cardiopulmonary Disease Associated with an Increase in Fine Particles Using Two-Stage Spatial Analytic Methods[a]

			Random Effects			
	Independent Observations		Independent Cities		Regional Adjustment	
Ecologic Covariate	Fine Particles	Covariate	Fine Particles	Covariate	Fine Particles	Covariate
Fine Particles Alone	1.30 (1.11–1.53)		1.38 (1.17–1.62)[c]		1.24 (1.01–1.52)	
Demographic Factors						
Population change	1.12 (0.99–1.26)	0.81 (0.74–0.88)	1.19 (1.00–1.43)	0.81 (0.71–0.93)	1.20 (0.95–1.51)	0.95 (0.79–1.15)
Whites	1.44 (1.28–1.61)	1.18 (1.08–1.29)	1.45 (1.23–1.70)[c]	1.15 (1.00–1.32)	1.29 (1.05–1.59)	1.10 (0.97–1.26)
Blacks	1.41 (1.24–1.59)	0.87 (0.78–0.98)	1.45 (1.22–1.72)[c]	0.88 (0.75–1.05)	1.28 (1.03–1.59)	0.92 (0.76–1.12)
Socioeconomic Factors						
Income	1.30 (1.17–1.44)	0.89 (0.81–0.97)	1.35 (1.15–1.59)[b]	0.91 (0.78–1.07)	1.21 (1.01–1.45)	0.81 (0.71–0.91)
Poverty	1.35 (1.21–1.50)	0.87 (0.78–0.98)	1.41 (1.19–1.66)[c]	0.91 (0.77–1.07)	1.21 (0.99–1.49)	1.12 (0.92–1.35)
Income disparity	1.36 (1.22–1.51)	0.82 (0.76–0.89)	1.40 (1.21–1.63)	0.82 (0.73–0.93)	1.25 (1.02–1.53)	0.93 (0.79–1.09)
Unemployment	1.25 (1.12–1.39)	1.15 (1.04–1.26)	1.33 (1.13–1.57)[b]	1.10 (0.96–1.27)	1.23 (1.01–1.50)	1.11 (0.98–1.27)
Education	1.27 (1.14–1.43)	0.96 (0.87–1.05)	1.37 (1.13–1.65)[b]	0.99 (0.85–1.15)	1.20 (0.99–1.46)	0.85 (0.71–1.00)
Health Services						
Physicians	1.30 (1.17–1.45)	0.86 (0.77–0.97)	1.36 (1.15–1.61)[b]	0.91 (0.76–1.10)	1.22 (0.98–1.52)	0.88 (0.74–1.05)
Hospital beds	1.29 (1.16–1.43)	1.13 (0.98–1.29)	1.37 (1.16–1.62)[b]	1.14 (0.94–1.38)	1.23 (0.98–1.53)	1.13 (0.93–1.37)
Climate						
Temperature	1.19 (1.07–1.33)	0.83 (0.77–0.90)	1.27 (1.09–1.47)	0.83 (0.74–0.93)	1.22 (1.00–1.50)	0.99 (0.80–1.24)
Temperature variation	1.14 (1.01–1.28)	1.22 (1.11–1.33)	1.21 (1.04–1.42)[b]	1.21 (1.06–1.38)	1.17 (0.94–1.45)	1.09 (0.92–1.28)
Relative humidity	1.20 (1.04–1.37)	1.07 (0.97–1.19)	1.26 (1.03–1.54)[c]	1.12 (0.95–1.33)	1.17 (0.90–1.53)	0.89 (0.66–1.21)
Relative humidity variation	1.27 (1.11–1.46)	0.90 (0.81–1.01)	1.33 (1.09–1.63)[c]	0.92 (0.77–1.11)	1.19 (0.92–1.53)	0.85 (0.66–1.10)
Physical Environment						
Altitude	1.26 (1.12–1.42)	1.08 (0.99–1.19)	1.30 (1.06–1.59)	1.00 (0.85–1.17)	1.18 (0.92–1.52)	1.03 (0.86–1.24)
Water hardness	1.27 (1.14–1.41)	1.20 (1.09–1.31)	1.34 (1.13–1.58)	1.19 (1.03–1.37)	1.22 (0.99–1.51)	1.17 (1.00–1.37)
Gaseous Copollutants						
CO	1.32 (1.19–1.47)	0.90 (0.83–0.98)	1.38 (1.17–1.64)	0.90 (0.78–1.05)	1.26 (1.02–1.56)	0.90 (0.77–1.06)
NO_2	1.39 (1.22–1.59)	0.91 (0.83–1.00)	1.51 (1.24–1.83)	0.91 (0.78–1.07)	1.34 (1.05–1.70)	0.87 (0.71–1.08)
O_3	1.31 (1.18–1.46)	0.94 (0.84–1.05)	1.38 (1.17–1.63)[b]	0.98 (0.83–1.17)	1.18 (0.97–1.43)	1.20 (1.02–1.41)
SO_2	1.17 (1.03–1.33)	1.47 (1.28–1.70)	1.25 (1.05–1.49)	1.40 (1.14–1.72)	1.23 (0.97–1.55)	1.26 (1.00–1.58)

[a] Effects were evaluated for the respective range of each covariate specified in Appendix G, which is available upon request from the Health Effects Institute. Data are RRs with 95% CIs.

[b] The residuals are spatially autocorrelated ($P < 0.05$) using the Moran I statistic.

[c] The residuals are spatially autocorrelated ($P < 0.10$) using the Moran I statistic.

CI: 0.91–1.28), whereas that for exposure to sulfur dioxide was 1.31 (95% CI: 1.12–1.28). Joint modeling produced a larger reduction in the sulfate-associated relative risk of mortality (1.20 to 1.08) than in the sulfur dioxide relative risk (1.35 to 1.31). The sulfate relative risk varied slightly less (8%) in terms of absolute amount [(1 – RR) × 100] (Independent Cities Model: RR = 1.13, 95% CI: 1.02–1.25; Regional Adjustment Model: RR = 1.10, 95% CI: 0.97–1.24; Filtered Both Sides Model: RR = 1.05, 95% CI: 0.97–1.14; Simultaneous Autoregressive Model: RR = 1.08, 95% CI: 0.91–1.28) than did the relative risk associated with sulfur dioxide (12%) (Independent Cities Model: RR = 1.27, 95% CI: 1.15–1.40; Regional Adjustment Model: RR = 1.23, 95% CI: 1.11–1.36; Filtered Both Sides Model: RR = 1.19, 95% CI: 1.10–1.29; Simultaneous Autoregressive Model: RR = 1.31, 95% CI: 1.12–1.54) when both pollutants were examined together.

Random Effects Cox Models

The original regression analyses of both the Six Cities Study and the ACS Study using the standard Cox model had been predicated on the assumption that the vital status of all study participants represented statistically independent

Table 49. Impact of Multiple Ecologic Covariates on the Relative Risks of Mortality from Cardiopulmonary Disease Associated with an Increase in Fine Particles Using Two-Stage Spatial Analytic Methods[a]

Spatial Analytic Model	Number of Cities	Relative Risk Calculated for Fine Particles Alone	Ecologic Covariates Incorporated into Adjusted Analyses	Relative Risk After Adjusting for Ecologic Covariates[b]	
				Fine Particles	SO_2
Demographic and Socioeconomic Factors					
Independent Observations	48	1.30 (1.17–1.45)	Population change, income, poverty, income disparity, unemployment, education	1.16 (1.00–1.35)	NC
Independent Cities	48	1.38 (1.16–1.63)		1.19 (0.98–1.45)	NC
Regional Adjustment	48	1.22 (0.97–1.52)		1.13 (0.91–1.40)	NC
Gaseous Copollutants					
Independent Observations	28	1.32 (1.16–1.50)	CO, NO_2, O_3, SO_2	1.22 (1.05–1.42)	1.45 (1.16–1.80)
Independent Cities	28	1.48 (1.22–1.80)		1.28 (1.05–1.57)	1.40 (1.13–1.73)
Regional Adjustment	28	1.40 (1.08–1.80)		1.26 (0.96–1.66)	1.21 (0.89–1.65)
25%[c]					
Independent Observations	32	1.27 (1.12–1.43)	Population change, whites, temperature variation, SO_2	1.18 (1.00–1.40)	1.23 (1.02–1.48)
Independent Cities	32	1.41 (1.17–1.69)	Population change, temperature variation, SO_2	1.10 (0.91–1.34)	1.27 (1.00–1.61)
Regional Adjustment	38	1.34 (1.07–1.70)	SO_2	1.23 (0.97–1.55)	1.26 (1.00–1.58)

[a] Effects were evaluated for the respective range of each covariate specified in Appendix G, which is available upon request from the Health Effects Institute.

[b] NC = Relative risks associated with SO_2 were not calculated in some multivariate models. Data are RRs with 95% CIs.

[c] Incorporated all ecologic covariates that, when analyzed individually in a bivariate model, were found to produce a change of 25% or more in the excess relative risk associated with fine particles.

outcomes. Because the life and death of each individual depends in a complex way on a number of health determinants, including characteristics of the city within which the subject resides, it was important to account in the analysis for individual heterogeneity as well as potential intracity correlation.

As we noted previously, the Reanalysis Team used two-stage models to address spatial clustering by city and region. We also employed two-stage spatial filtering methods to take into account more complex spatial patterns in the data. In addition to the two-stage random effects methods used to address spatial autocorrelation at different levels, we developed powerful new methods of incorporating random effects into the Cox regression model (Appendix I). Specifically, we considered a Cox proportional-hazards model with a random effect that represents the unique characteristics of each city. This approach avoids the approximations inherent in the two-stage random effects models by estimating the regression parameters and random effects within a single integrated estimation framework.

The random effects Cox model assumes that, given the city-specific random effects, the hazard functions for individuals are conditionally independent, with the hazard function for individual j from city i given by

$$h_{ij(s)}(t) = h_s(t)u_i \exp(\beta^T \mathbf{x}_{ij(s)})$$

at time t, where the subscript s denotes different strata within the age-stratified and gender-stratified cohort. The city-specific random effects u_i are assumed to follow flexible Tweedie distributions with unit mean and variance τ^2. The regression vector β reflects the effects of the covariate vector $\mathbf{x}_{ij(s)}$ on the baseline hazard $h_s(t)$ in each stratum s. Further details of the statistical methods used to fit the random effects Cox model are given in Appendix I.

Although the opportunity to characterize intercity variation is limited by the number of cities in the Six Cities Study, this is not the case in the ACS Study, which involved 151 cities in the sulfate cohort and 50 cities in the fine particle cohort. Table 50 shows the relative risks of mortality from all causes for exposure to both fine particles and sulfate based on our random effects Cox model fit to

Table 50. Relative Risks of Mortality from All Causes Associated with Sulfate and Sulfur Dioxide in the Reanalysis of the ACS Study Comparing Single and Multiple Pollutants in a Standard Cox Model, a Two-Stage Model, and a Random Effects Cox Model

Pollutant	Number of Cities	Standard Cox Model[a]	Two-Stage Independent Cities Model[b]		Random Effects Cox Model[b]	
		RR (95% CI)	RR (95% CI)	τ^2	RR (95% CI)	τ^2
ACS Fine Particle Cohort						
Single-pollutant model						
$PM_{2.5}$	52	1.18 (1.10–1.27)	1.31 (1.14–1.51)	0.0067	1.31 (1.14–1.49)	0.0056
SO_2	38	1.53 (1.40–1.68)	1.55 (1.32–1.81)	0.0036	1.50 (1.29–1.74)	0.0034
Multiple-pollutant model						
$PM_{2.5}$	38	1.02 (0.94–1.12)	1.14 (0.98–1.33)	0.0041	1.13 (0.97–1.31)	0.0034
SO_2	38	1.52 (1.37–1.68)	1.44 (1.20–1.72)		1.40 (1.18–1.66)	
ACS Sulfate Cohort						
Single-pollutant model						
SO_4^{2-}	38	1.16 (1.10–1.23)	1.25 (1.13–1.37)	0.0050	1.22 (1.12–1.34)	0.0040
SO_2	151	1.33 (1.25–1.41)	1.33 (1.22–1.46)	0.0028	1.31 (1.21–1.43)	0.0023
Multiple-pollutant model						
SO_4^{2-}	113	1.05 (0.98–1.12)	1.13 (1.02–1.25)	0.0029	1.12 (1.02–1.23)	0.0023
SO_2	113	1.30 (1.22–1.39)	1.27 (1.15–1.39)		1.25 (1.14–1.37)	

[a] This model had the same covariate specifications used by the Original Investigators.

[b] τ^2 is the dispersion parameter.

the ACS data, along with the relative risks based on both the standard Cox model used by the Original Investigators and our two-stage Independent Cities Model. In the fine particle cohort, the relative risk of all-cause mortality associated with fine particle air pollution based on the random effects Cox model (RR = 1.31, 95% CI: 1.14–1.49) is virtually identical to that based on the two-stage approach (RR = 1.31, 95% CI: 1.14–1.51), which confirms the validity of the approximate but computationally simpler two-stage random effects approach. However, these relative risks are notably greater than that based on the standard Cox model (RR = 1.18, 95% CI: 1.10–1.27), which does not acknowledge intracity correlation. The overdispersion parameter $\tau^2 = 0.0067$ based on the two-stage method also is comparable to the value of $\tau^2 = 0.0056$ achieved by the full random effects Cox model.

When we calculated the relative risk of mortality associated with exposure to fine particles using the standard Cox model, it was notably lower than that based on the random effects Cox model. This occurred because, as in two-stage regression, the random effects Cox model took into account heterogeneity among cities as measured by the overdispersion parameter τ^2. The inclusion of this between-city variation in the weighting scheme gave less weight to larger cities, which in this case resulted in an increased relative risk.

We conducted a similar analysis on the fine particle cohort using sulfur dioxide in place of fine particles. As was the case with fine particles, the relative risk of all-cause mortality associated with exposure to sulfur dioxide obtained with the random effects Cox model (RR = 1.50, 95% CI: 1.29–1.74) was similar to that obtained with the two-stage Independent Cities Model (RR = 1.55, 95% CI: 1.32–1.81). In this case, however, we obtained a similar relative risk from the standard Cox model (RR = 1.53, 95% CI: 1.40–1.68) as well. This occurred because the heterogeneity among cities was lower for sulfur dioxide ($\tau^2 = 0.0036$ under the two-stage model) than for sulfate ($\tau^2 = 0.0067$). Because sulfur dioxide exhibited less overdispersion, the city-specific weights used in the two-stage Independent Cities Model were similar to those in the two-stage Independent Observations Model, which, as shown previously, produced relative risks similar to those from the standard Cox model.

When we included both fine particles and sulfur dioxide as predictors of mortality in the same random effects Cox model, we reduced the relative risk of mortality associated with fine particles from 1.31 to 1.13. Note,

however, that the relative risk of mortality from all causes associated with exposure to fine particles remained elevated even after adjustment for the effects of sulfur dioxide.

Because the overdispersion parameter $\tau^2 = 0.0034$ in the single-pollutant sulfur dioxide model was identical to that in the model that included both sulfur dioxide and fine particles, fine particles did not appear to predict variation in all-cause mortality among cities beyond estimates provided by sulfur dioxide alone.

The sulfate cohort exhibited patterns similar to those just described for the fine particle cohort. Specifically, the two-stage random effects Independent Cities and random effects Cox models led to similar results, with exposure to sulfur dioxide in air pollution explaining more of the variation in mortality than did exposure to sulfate. However, the effects of sulfate remained significant even after we adjusted for sulfur dioxide under the two-stage Independent Cities Model and the random effects Cox model.

Effects of Sulfate and Sulfur Dioxide by Region

Our spatially filtered analysis compares high (low) mortality rates for a spatially local area (600 km in radius) to high (low) air pollution levels in the corresponding area. Thus, we remove broad spatial patterns before we link the variables together. This type of analysis is conducted to remove the possibility that coincidental broad spatial patterns in both variables will influence the associations between mortality and air pollution. This type of adjustment is of interest because there may be important risk factors for mortality that we have not accounted for in our analysis.

Spatial filtering can adjust for risk factors, such as dietary habits, that aggregate at a broad spatial level. When we compared relative risks obtained from the Independent Cities Model and the Spatial Filtering Model, we could estimate how much of the air pollution effect on mortality was attributable to broader spatial patterns and where the effect existed on a more local level. Reduction in the relative risk associated with sulfate exposure from 1.25 (Independent Cities Model) to 1.19 (Regional Adjustment Model) to 1.09 (Spatial Filtering Model) suggests that most of the association between sulfate and mortality resulted from the spatial coincidence of these variables on a relatively large scale.

The seven regions selected for our Regional Adjustment Model were similar in size to the area needed to remove spatial autocorrelation in the spatial filtering analysis. The main difference between these two types of analyses is that the latter removes the broad spatial patterns in sulfate before sulfate values are linked with the spatially filtered mortality rates. There may be some concordance in space between mortality rates and sulfate exposure that is not accounted for by the Regional Adjustment Model. We examined this possibility by conducting separate analyses by region.

Four of the regions originally examined had too few cities with either sulfate or sulfur dioxide data (the Northwest had only 16 cities with sulfate data and 10 cities with

Table 51. Relative Risks of Mortality from All Causes Associated with Sulfate and Sulfur Dioxide in the Reanalysis of the ACS Study by Region Using a Two-Stage Model[a]

	Sulfate				Sulfur Dioxide			
Region	Number of Cities	Mean ($\mu g/m^3$)	Range ($\mu g/m^3$)	RR (95% CI)[b]	Number of Cities	Mean ($\mu g/m^3$)	Range ($\mu g/m^3$)	RR (95% CI)[b]
Single-Pollutant Model								
Northeast	41	11.5	12.8	1.14 (0.93–1.40)	36	11.8	15.6	1.20 (1.00–1.45)
Industrial Midwest	34	13.3	14.6	1.29 (1.07–1.55)	30	11.1	25.7	1.24 (1.11–1.38)
Southeast	30	11.6	11.9	1.25 (1.01–1.54)	19	6.6	13.4	1.29 (0.98–1.70)
West[c]	46	7.1	11.0	0.91 (0.71–1.17)	28	5.9	16.5	1.30 (1.00–1.67)
Two-Pollutant Model								
Northeast	36			1.03 (0.85–1.24)	36			1.19 (0.98–1.45)
Industrial Midwest	30			1.09 (0.88–1.35)	30			1.19 (1.04–1.38)
Southeast	19			1.30 (0.99–1.70)	19			1.10 (0.82–1.48)
West[c]	28			0.91 (0.72–1.16)	28			1.31 (1.01–1.69)

[a] The Random Effects Cox Model assumes cities are statistically independent.

[b] Relative risks were calculated at the range for each pollutant across the entire study dataset.

[c] Data from the Northwest, Southern California, Southwest, and Upper Midwest were combined to form the West region.

sulfur dioxide data; Southern California, 6 and 5; Southwest, 10 and 6; and Upper Midwest, 14 and 7, respectively). Data for these four regions were combined to form a West region. The Independent Cities Model, which assumed that the cities within each region were statistically independent, was run for sulfate alone, sulfur dioxide alone, and both pollutants together for each of the four regions (Northeast, Industrial Midwest, Southeast, and West).

Both sulfate and sulfur dioxide levels were high in the Northeast and Industrial Midwest, with lower concentrations in the West (Table 51). Sulfate was correlated weakly with sulfur dioxide in the Northeast ($r = 0.18$) and West ($r = 0.17$), correlated moderately in the Southeast ($r = 0.44$), and correlated highly in the Industrial Midwest ($r = 0.69$).

Three of the four regions exhibited positive associations between exposure to sulfate and deaths from all causes. We observed an inverse association in the West where mean sulfate levels are lowest. We observed positive associations between mortality and exposure to sulfur dioxide in all four regions, with the largest relative risk of mortality from all causes found in the West. The relative risk of mortality associated with sulfur dioxide was larger than that associated with sulfate in three regions, with only the Southeast region displaying a greater sulfate effect based on the two-pollutant model specifications. The two-pollutant model, with its strong negative association with sulfate in the West and corresponding strong positive association with sulfur dioxide, suggests why sulfur dioxide accounts for much of the sulfate effect on mortality when all cities are examined together. The sulfur dioxide effect was insensitive to adjustment for sulfate in all four regions; the sulfate effect, however, changed considerably in all but one of the four regions (the Southeast) after adjustment for sulfur dioxide.

DISCUSSION

The association between fine particle pollution in ambient air and cardiorespiratory morbidity and mortality has been explored in a number of epidemiologic investigations; both time-series and cohort studies have shown positive associations between ambient fine particles and mortality. The Six Cities Study and the ACS Study provided important information on this association and were the basis for the promulgation of the first US National Ambient Air Quality Standard (NAAQS) for fine particles. Positive associations between ambient fine particles and mortality had been demonstrated in the original analyses of these two large-scale cohort studies. For instance, in the Six Cities Study, the adjusted mortality rate ratio for the most-polluted city compared to the least-polluted city was 1.26 (95% CI: 1.08–1.47). In the ACS Study, the all-cause mortality risk ratio for the most-polluted city compared with the least-polluted city was 1.17 (95% CI: 1.09–1.26).

DESIGN OF THE ORIGINAL STUDIES

Although these two studies produced comparable results, they differed markedly in design. The Six Cities Study was a prospective cohort study, with subjects recruited between 1974 and 1977 from six cities (Watertown MA, Harriman TN, and Steubenville OH, St Louis MO, Portage WI, and Topeka KA) and followed for up to 16 years. The cities, located in the Midwest and Northeast United States, had been chosen to represent a gradient in ambient air pollution. The original cohort comprised 8,111 white adults, 25 to 74 years of age. All subjects had completed a standardized questionnaire that elicited information concerning age, gender, weight, height, education level, complete smoking history, occupational exposures, and medical history.

The Six Cities Study had a number of strengths, including random selection of study subjects; reasonably high participation rates; personal interviews with all respondents at the time of enrollment; subsequent follow-up at intervals of 3, 6, and 12 years; and pulmonary lung function testing using appropriate spirometric techniques. Exposure monitoring was conducted largely by ambient air pollution monitors developed and operated by the Original Investigators at Harvard University, although data from the US EPA's AIRS database also were used. The baseline questionnaire administered at the time of enrollment was extensive and included items on age, gender, weight, height, education level, complete smoking history, occupational exposures, and health status. Residence histories before and after enrollment were recorded, which permitted direct assessment of residential mobility of the study subjects.

The ACS Study drew on a larger cohort from the CPS-II, which involved approximately 1.2 million individuals. The cohort assembled by the Original Investigators included 552,138 persons in 154 United States cities located in all 50 states. Participants were at least 30 years old and were members of households with at least one individual 45 years old or older. Because the CPS-II had not been designed expressly to address the relation between ambient fine particle concentrations and mortality, the Original Investigators did not develop questionnaire items specific to this purpose. Nonetheless, the ACS

questionnaire included a rich set of items providing information on health status, demographic characteristics, smoking history, alcohol use, and occupational exposures to pollutants.

The ACS Study cohort was recruited in 1982, with mortality follow-up through 1989. Vital status was ascertained by personal inquiries by volunteers in September 1984, 1986, and 1988. Automated linkage using the National Death Index (NDI) maintained by the National Center for Health Statistics was used to extend the follow-up to December 31, 1989. Ambient levels of fine particles and sulfate were obtained from two sources. Mean concentrations of sulfate air pollution for 1980 were obtained from the EPA's NAD and the EPA's IPMN. Median fine particle concentrations for 1979 through 1983 were calculated from the EPA's IPMN using dichotomous samplers. Whereas sulfate air pollution data were available for 151 United States cities, fine particle data were taken from previously published data for only 50 United States cities.

The Six Cities Study and the ACS Study possess complementary strengths and limitations. Although the Six Cities Study had been designed specifically to test the hypothesis that long-term exposure to fine particle air pollution is associated with increased mortality rates, the study involved only six cities within a limited geographic region of the United States. Because only one pollution monitor was positioned in each city, all individuals within the city were assigned the same level of exposure for each of the pollutants considered. Thus, although a large number of individual covariates had been recorded for each of the 8,111 subjects in the Six Cities Study, the limited nature of the pollution monitoring reduced the effective number of data points in the exposure-response gradient to six and uncertainty in estimating the citywide averages effectively reduced the number of data points even further. Further adjustment for the effects of other ecologic covariates in the Six Cities Study was not practical because of the limited number of degrees of freedom (at most 6 *df*) for further analyses. The ACS Study, which involved 154 cities with a wide range of pollutant concentration profiles, was not seriously affected by this limitation.

The different nature of the two studies provided the Reanalysis Team with opportunities to explore the sensitivity of the original findings to alternative analytic approaches and to incorporate additional data not explicitly considered in the original publications.

LIMITATIONS OF THE ORIGINAL STUDIES

The Six Cities Study and the ACS Study had certain limitations, including the inability to strictly characterize the long-term exposure of study participants to fine particles (Vedal 1997). In both studies, exposure to ambient air pollutants necessarily had been gauged using fixed-site ambient monitors, as personal dosimetry for such large cohorts would have been both impractical and prohibitively expensive. Abbey and colleagues (1999) also relied on fixed-site ambient monitors maintained by the California Air Resources Board in the Seventh-Day Adventist Study. The use of such area monitors leads to some degree of exposure measurement error (Clayton et al 1993; Kotchmar et al 1987; Leaderer et al 1999), the consequences of which have been discussed below. Nonetheless, fixed-site monitors are widely used in large-scale studies of air pollution and population health, including informative time-series studies of the association between ambient fine particles and morbidity and mortality in the general population (Burnett et al 1995, 1998; Samet et al 2000).

Other potential limitations of the Six Cities Study and the ACS Study cited by other investigators include inadequate control of age and sedentary lifestyle (Moolgavkar and Luebeck 1996), and insufficient control of cigarette smoking, both active and passive (US EPA 1996). Our analytic plan (Krewski et al 1998) called for our reanalysis to exert the maximal control possible for potential confounding due to these and other covariates on which information was available. For the ACS Study, we also assembled and used a series of additional ecologic covariates that represented potential determinants of population health in a further attempt to control for confounding.

Gamble (1998) prepared a detailed critique of the Six Cities Study and the ACS Study. He focused particularly on the ecologic nature of the exposure measurements resulting from the use of fixed-site ambient monitors, as noted previously. Gamble suggested that lung function (as measured by FEV_1) and sedentary lifestyle also could be important confounding variables, and that there could be residual confounding from a failure to consider nonlinear effects of alcohol consumption and body weight. To address these latter concerns, the Reanalysis Team accounted for possible nonlinear effects of these covariates and included spirometric pulmonary function measurements in the reanalysis.

Information on population mobility was lacking in both studies. To evaluate population mobility in the Six Cities Study, we coded the residence histories that had been recorded, but not examined in detail, by the Original Investigators. Because residential addresses were available only at the time of recruitment into the cohort in 1980 in the ACS Study (and at the time of death for those subjects who died during the period of follow-up), we could not evaluate mobility in the ACS Study cohort.

REPLICATION AND VALIDATION

The results of the Reanalysis Team's replication and validation of the original findings are presented in Part I of the Investigators' Report. We validated those findings by comprehensively auditing all variables that had been used in the original published analyses, comparing the original information (eg, from questionnaires and air pollution monitors) with the data in the analysis files. Our replication involved duplicating the selection process that had defined the subcohorts in the original analysis and replicating the original numerical results using the same analytic methods reported by the Original Investigators. Although we identified some discrepancies in questionnaire-based items and vital status, we found the data from both studies to be of generally high quality. We could not trace all the original air pollution data from the Six Cities Study; nonetheless, there was sufficient evidence to confirm the integrity of the long-term average fine particle levels for each of the six cities in the study. Although we noted some discrepancies in the selection and follow-up of the subcohorts used in both the Six Cities Study and the ACS Study, these discrepancies did not dramatically affect the risk estimates in either study. The Reanalysis Team, using the same data and analytic techniques that had been used by the Original Investigators, concluded that the original findings in both studies were substantiated.

DATA QUALITY AUDIT

Part I of the reanalysis included a detailed audit of all variables that had been used by the Original Investigators. In keeping with our intent to audit all variables involved in the reanalysis, we subjected individual-level variables used for the first time in Part II to the same rigorous audit standards. As in Part I, we found few errors in most variables, although we did find a number of errors in the occupational coding that had been assigned in the ACS Study. Consequently, our ability to control for occupational confounding in the ACS Study was limited by the quality of the occupational data.

SENSITIVITY ANALYSES

The Reanalysis Team conducted a number of different sensitivity analyses to further explore the associations between mortality and fine particles or sulfate. First, we explored the sensitivity of the Original Investigators' risk estimates to the inclusion of additional variables in the Cox proportional-hazards model, and to different ways of characterizing variables such as education, which we considered explicitly in the Original Models. These analyses had two related but distinct objectives: to identify potential confounding variables of the association between mortality and fine particles, and to identify variables that modify the effect of fine particle air pollution.

Using two new aggregate indices, we also investigated in detail the possibility of confounding due to occupational exposures. The first index provided a seven-level ordinal measure of the overall dirtiness of specific jobs and occupations of the study subject; the second provided a binary indicator of ever/never having been exposed in the workplace to agents accepted as being associated with increased lung cancer risk. Members of the Reanalysis Team who have extensive experience in occupational exposure assessment developed these aggregate occupational exposure indices on the basis of occupational and industrial codes that had been assigned by the Original Investigators.

In the Six Cities Study, the availability of additional data on study subjects at 3, 6, and 12 years after the collection of baseline data upon enrollment permitted us to assess changes in key covariates, such as tobacco consumption and BMI, over time. Likewise, we were able to assess the impact of population mobility on estimates of risk because detailed residence histories had been included in this study. The ACS Study, which involved 154 MSAs from across the United States, allowed us to assess changes in risk associated with a number of auxiliary sociodemographic and environmental variables derived from publicly available data sources. We used random effects methods and flexible nonparametric risk models to assess variation in mortality rates among cities.

We outlined these sensitivity analyses in the Analytic Plan prepared before we began the reanalysis (Krewski et al 1998). In addition to the planned analyses, we applied modeling techniques that controlled for spatial autocorrelation in measures of fine particle air pollution and other ecologic covariates.

Two of our planned analyses were not attempted. Specifically, we didn't make comparisons by race because of the small number of minority subjects in both cohorts, and we didn't perform the proposed exploration of critical exposure-time windows (ie, the period of exposure most strongly associated with mortality) in the Six Cities Study for several practical reasons. First, the residential mobility information needed to accurately characterize exposure for the period before enrollment is incomplete. (The Reanalysis Team did, however, construct postenrollment residence histories on the basis of information from the follow-up questionnaires and postcards.) Second, postenrollment mobility was limited; only 18.5% of the study subjects left the original city of residence during the follow-up period. And third, historical records of fine particle levels are not

available for these cities before 1980, which precluded evaluating fine particle exposures in the early years of the Six Cities Study. Despite these limitations, the Six Cities Study does present an opportunity to evaluate the effects of changes in exposure over time, including possibly important exposure-time windows. This analysis would be most informative if exposures in all residences outside the six cities were assessed by spatial interpolation methods, a complicated task that is outside the scope of the reanalysis.

The results of the sensitivity analyses conducted in Part II of the reanalysis are summarized below.

Alternative Risk Models

We considered an extensive series of alternative risk models. The Reanalysis Team found little evidence that questionnaire variables had led to confounding in either study, thereby strengthening the conclusion that the observed association between fine particle air pollution and mortality was not the result of a critical covariate that had been neglected by the Original Investigators. Although this analysis is reassuring, it does not rule out the possibility of confounding by unmeasured covariates. The Reanalysis Team also found that the risk estimates in both studies were not sensitive to the manner in which individual covariates were characterized.

The Reanalysis Team tested the goodness of fit of the Cox proportional-hazards model in the Six Cities Study. Although the model did not demonstrate a significant lack of fit overall, there was some evidence that the effects of both fine particles and sulfate on mortality varied with time.

Controlling for the Effects of Age

Because the original study outcomes were strongly dependent on age, and because of the possibility of differing age structures across the cities represented in the two studies, we attempted to account more precisely for the effects of age. One method is to use age as the time axis, rather than calendar year, in the proportional-hazards model (Breslow and Day 1987). For most causes of death, these two methods of controlling for the effects of age produced comparable results.

Controlling for Other Covariates

In the original analyses, the data had been stratified by age (5-year categories) and gender. The ACS Study also had been stratified by race (white, black, and other). The following covariates had been included in the original analyses for the Six Cities Study: an indicator variable for current- or former-smokers, number of pack-years for current-smokers, number of pack-years for former-smokers, an indicator variable for less than high school education, and BMI (weight in kilograms divided by the square of height in meters; also referred to as the Quetelet Index). For the ACS Study, the statistical adjustments had included an indicator variable for current-smokers; an indicator variable for pipe- and/or cigar-smoker only; number of years smoked for current-smokers; number of cigarettes per day for current-smokers; number of years smoked for former-smokers; number of cigarettes per day for former-smokers; number of hours per day exposed to environmental tobacco smoke; BMI; number of drinks per day of alcoholic beverages; an indicator variable for less than high school education; and an indicator variable for regular occupational exposure to any asbestos, chemicals/acids/solvents, coal or stone dust, coal tar/pitch/asphalt, diesel engine exhaust, or formaldehyde.

We extended these statistical models by incorporating a wide range of individual covariates that included finer levels of adjustment, adding quadratic terms for some variables and considering gender differences (statistical interactions) in the effects of these variables. We examined as separate subgroups those individuals who, at the beginning of the study period, reported selected diseases (high blood pressure, heart disease, stroke, diabetes, chronic bronchitis and emphysema, asthma, or any cancer). In addition, we examined the potentially confounding effect of physical exercise on the relation between air pollution and mortality by including the self-reported amount of physical exercise (none or some, moderate, or heavy) at time of enrollment in the model. Because we postulated that illness causes stricken individuals to exercise less than healthy persons, we examined the group that had not reported having the selected diseases in order to minimize the potential that level of exercise was a variable in the causal pathway. Again, the results were essentially identical to those from the other models. We concluded that finer levels of control for these measured covariates did not alter the original findings of an association between air pollution and mortality.

Influence of City

In studies involving multiple cities, the overall results may be unduly influenced by a single city. This is particularly likely in the Six Cities Study, which involved a small number of communities. In an attempt to identify strong leverage points, we estimated the effect that each city had on the estimated hazard ratios by excluding in turn each city from the analysis (comparable to deletion regression diagnostics). We found that the results were not influenced by the exclusion of any of the six cities. This means that the form of the exposure-response pattern, as well as the

estimated slope, was not seriously influenced by cities with less pollution (Portage WI) or more pollution (Steubenville OH). We conducted a similar regression analysis for the ACS Study and found that no single city exerted an undue influence on the association between air pollution and mortality.

Because the original air pollution monitoring records in the ACS Study were unavailable for audit, the Reanalysis Team constructed alternative air pollution data on the basis of monitoring data accumulated by the US EPA. Although the city-specific fine particle air pollution levels that had been assembled by the Original Investigators correlated highly with those developed by the Reanalysis Team, there were notable differences in fine particle levels for Denver. However, inclusion or exclusion of Denver from the reanalysis had no appreciable effect on the overall mortality risk ratios.

Identification of Sensitive Subpopulations

The Reanalysis Team examined the changes in relative risk estimates associated with air pollution for specific subgroups of the study populations (statistical interactions), conducting separate analyses for well-defined categories of each of the following variables: age at enrollment; gender; educational attainment; marital status; smoking status; diseases reported at time of enrollment; amount of time lived in neighborhood at time of enrollment; self-reported occupational exposure to toxic air (dust, gases, and fumes); and our own lung carcinogen occupational dirtiness indices.

In the ACS Study, we found no important differences in relative risk of mortality by gender or age at enrollment. Although the estimates of the relative risk associated with air pollution differed for other variables, the 95% confidence intervals overlapped in all instances, so that the differences could be formally explained by chance alone.

The Reanalysis Team did find strong evidence of effect modification for some variables. Education notably modified the air pollution–mortality association (for both fine particles and sulfate); individuals who did not complete high school had the highest relative risks of mortality. Conversely, individuals who completed high school did not appear to have had increased risk. The Reanalysis Team concludes that this modifying effect is not necessarily attributable to education per se, but could indicate that education is a marker for a more complex set of socioeconomic variables that impact upon the level of risk.

Comparison of Results Between Studies

Estimates of risk of mortality associated with exposure to fine particles and sulfate were insensitive to the set of covariates included in the risk model (the Original, Full, and Extended Model specifications yielded almost identical risks for the Six Cities Study and the ACS Study). In both studies, we obtained similar relative risks whether we used calendar year or age as the time axis. Also in both studies, cardiovascular disease mortality had the highest relative risk associated with exposure to air pollution. We found no associations between air pollution and death from respiratory disease in either study. In the Six Cities Study, the relative risks of other causes of death were similar to those for death from all causes, whereas the relative risks of other causes of death were much lower in the ACS Study. In the ACS Study, we observed slightly larger relative risks for the other cancers group (RR from 1.08 to 1.14) compared with those in the Six Cities Study (RR 1.03 to 1.04).

Although the air pollution effect was less among married persons in the ACS Study, the relative risks of mortality in the Six Cities Study were independent of marital status. Gender did not modify the mortality effect of fine particle air pollution in the ACS Study, but did so in the Six Cities Study.

The relative risks of mortality associated with an increase in exposure to fine particles or sulfate, by underlying cause of death and educational attainment, are shown in Table 52. Although relative risks clearly declined with increasing educational attainment for all causes of death examined in the ACS Study, this pattern was not as consistent in the Six Cities Study.

Flexible Risk Models

Under the Cox proportional-hazards regression model, a fixed increment in ambient pollutant levels has the same multiplicative effect on the mortality rate at any point in time, so that the hazard functions for mortality at two levels of pollution are proportional. In addition, this model assumes that the relative increase in mortality is described by a specific parametric form; specifically, that the logarithm of the hazard rate is a linear function of the covariates.

To evaluate the applicability of this model to the Six Cities Study, the Reanalysis Team considered more flexible models, based on regression spline generalizations of the Cox model, to describe the relation between mortality and fine particles and sulfate. This flexible modeling approach indicated that the linearity assumption implicit in the Cox model was appropriate for fine particles. However, there was some evidence of departure from linearity at both low and high sulfate concentrations. Consistent with the quadratic relation between BMI and mortality in our Extended Model for both studies, the flexible modeling approach suggested a U-shaped relation between BMI and mortality.

Table 52. Relative Risks of Mortality by Cause of Death Associated with an Increase in Fine Particles or Sulfate by Education Level in the Reanalysis of the Six Cities and ACS Studies[a]

	ACS Study			Six Cities Study		
Cause of Death	Less Than High School [11%]	High School [30%]	More Than High School [59%]	Less Than High School [28%]	High School [38%]	More Than High School [34%]
Fine Particles						
All causes	1.35 (1.17–1.56)	1.23 (1.07–1.40)	1.06 (0.95–1.17)	1.45 (1.13–1.85)	1.30 (0.98–1.73)	0.97 (0.71–1.34)
Cardiopulmonary disease	1.47 (1.21–1.78)	1.35 (1.11–1.64)	1.14 (0.98–1.34)	1.28 (0.92–1.77)	1.42 (0.98–2.08)	1.40 (0.88–2.23)
Cardiovascular disease	1.47 (1.19–1.82)	1.39 (1.13–1.72)	1.24 (1.05–1.47)	1.31 (0.92–1.87)	1.63 (1.10–2.42)	1.37 (0.84–2.22)
Respiratory disease	1.36 (0.80–2.32)	1.16 (0.69–1.95)	0.65 (0.42–1.02)	0.97 (0.38–2.46)	0.36 (0.09–1.39)	1.80 (0.26–12.35)
Lung cancer	1.41 (0.87–2.29)	1.39 (0.90–2.15)	0.66 (0.46–0.95)	2.69 (1.09–6.60)	0.50 (0.11–2.22)	1.08 (0.33–3.58)
Other cancers	1.20 (0.87–1.66)	1.12 (0.87–1.43)	1.14 (0.94–1.38)	1.33 (0.75–2.37)	1.48 (0.77–2.83)	0.53 (0.25–1.09)
Other causes	1.12 (0.76–1.64)	1.00 (0.71–1.41)	0.95 (0.73–1.24)	1.76 (0.93–3.33)	0.65 (0.29–1.44)	0.69 (0.31–1.55)
Sulfate						
All causes	1.27 (1.13–1.42)	1.20 (1.08–1.33)	1.05 (0.96–1.14)	1.47 (1.14–1.89)	1.30 (0.97–1.73)	0.99 (0.72–1.36)
Cardiopulmonary disease	1.39 (1.20–1.62)	1.31 (1.13–1.53)	1.11 (0.98–1.25)	1.28 (0.91–1.79)	1.38 (0.94–2.02)	1.42 (0.90–2.24)
Cardiovascular disease	1.44 (1.23–1.69)	1.42 (1.21–1.67)	1.19 (1.05–1.36)	1.30 (0.90–1.87)	1.59 (1.06–2.37)	1.40 (0.87–2.26)
Respiratory disease	1.11 (0.74–1.66)	0.78 (0.52–1.18)	0.66 (0.47–0.93)	1.05 (0.40–2.72)	0.29 (0.07–1.24)	1.73 (0.26–11.38)
Lung cancer	1.49 (1.02–2.18)	1.39 (0.99–1.95)	1.19 (0.89–1.59)	2.82 (1.15–6.90)	0.51 (0.11–2.25)	0.91 (0.27–3.02)
Other cancers	0.97 (0.76–1.24)	1.28 (1.06–1.54)	1.04 (0.89–1.21)	1.44 (0.80–2.60)	1.56 (0.81–2.99)	0.59 (0.29–1.22)
Other causes	1.16 (0.86–1.56)	0.71 (0.55–0.94)	0.84 (0.68–1.04)	1.66 (0.86–3.19)	0.64 (0.28–1.44)	0.67 (0.30–1.50)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the Six Cities Study, this difference for fine particles was 18.6 $\mu g/m^3$ and for sulfate it was 8.0 $\mu g/m^3$; in the ACS Study, this difference for fine particles was 24.5 $\mu g/m^3$ and for sulfate it was 19.9 $\mu g/m^3$. Time axis was calendar year. Percentage of sample in educational group is given in square brackets. Data are RRs with 95% CIs.

Although the Cox proportional-hazards assumption did appear to be appropriate for most of the study period, there was some evidence that the effects of both fine particles and sulfate varied somewhat with time. (Recall that tests for lack of fit of the Cox proportional-hazards regression model also provided some evidence that the effects of fine particles and sulfate may vary over time.) The pattern of time dependency suggests that the multiplicative effect of recent exposures on hazard rates may be associated more strongly with mortality than are exposures that occurred many years before death.

Analyses of the ACS Study cohort did not identify a consistent pattern in changes over time of the impact of either fine particles or sulfate on mortality. However, flexible analyses of the ACS Study data yielded evidence of nonlinear exposure-response relations for both fine particles and sulfate, with sulfate demonstrating a comparatively shallow exposure-response relation at concentrations below 10–15 $\mu g/m^3$. Flexible analyses of the ACS Study data also demonstrated, as in the Six Cities Study, a nonlinear (U-shaped) relation of BMI with mortality.

Occupational Confounding

Although occupational exposures had been considered to some extent by the Original Investigators, the original analysis had been restricted in the Six Cities Study to self-reported occupational exposure to dust and fumes and in the ACS Study to a selected number of toxic air pollutants. The Reanalysis Team was concerned about the possibility of occupational confounding. For example, individuals in cities with high air pollution levels might tend to work in jobs that incurred high exposure to other agents associated with increased mortality. Consequently, our reanalysis used additional information on occupational exposures derived from the occupational histories available in both studies.

The Reanalysis Team developed and applied two aggregate indices of occupational exposures to the Six Cities Study and the ACS Study. The first index provided a mea-

sure of the overall dirtiness of the environments in which the subjects had worked. The second index reflected occupational exposure to accepted lung carcinogens.

The dirtiness index can be conceptualized as a variable that encompasses and integrates all exposures in the workplace, including exposure to pollutants that are pathogenic. But it can also be conceptualized as a variable that captures an aspect of social class that is correlated with, but distinct from, educational attainment—the other social class variable that had been used by the Original Investigators. The lung carcinogen index is a binary indicator variable that reflects whether or not a subject (ever/never) has been exposed occupationally to agents that have been identified as increasing risk of lung cancer.

We emphasize that there are four limiting factors associated with using the new indices to control confounding. First, the occupational information collected from study subjects did not represent detailed lifetime work histories. Second, the validity of the occupation coding has not been established in relation to the actual jobs and occupations held; because the Six Cities Study used a more detailed occupation coding system than did the ACS Study, there is greater potential in the former for valid attribution of both the dirtiness index and the lung carcinogen index. Third, the indices themselves are crude simplifications of complex exposure circumstances. In a sense, the indices are ecologic variables that establish an individual's presence within a potentially hazardous environment but do not measure individual exposure. Fourth, the two indices constructed focus on the dirtiness of jobs and subjects ever/never having been exposed to known lung carcinogens. It is recognized that the dirtiness index may be more useful in the case of deaths from respiratory conditions than deaths from cardiovascular disease. The lung carcinogen index is specifically designed to control for potential confounding by exposure in the workplace to agents known to increase lung cancer risk.

The impact of these limitations is to lessen the ability of the analyses to adequately adjust for potential confounding variables; that is, whatever bias in the original results might be due to confounding by occupational exposure would be diminished, but not necessarily eliminated, in our reanalyses. Nevertheless, we believe that this approach to controlling occupational confounding is an improvement over the original analyses. The new indices appeared to perform their intended function in that they were correlated with other variables in an expected way. The Six Cities Study had higher dirtiness scores, and higher prevalence of occupational exposure to carcinogens, than the ACS Study, compatible with what is known about the respective study populations. The dirtiness index was not correlated with air pollution in the ACS Study, but it was in the Six Cities Study. Although the dirtiness index was not a risk factor for mortality in the ACS Study, it was in the Six Cities Study.

The inclusion of these new variables had almost no impact on the relative risks of air pollution for cardiopulmonary mortality and mortality from all causes. In the ACS Study, we found excess risks for lung cancer from exposure to sulfate pollution but not fine particle pollution; lung cancer risks exhibited little change after adjusting for occupation. In the Six Cities Study, we found a nonsignificant excess in lung cancer risk related to fine particle air pollution, although this risk was attenuated considerably when the occupational confounders were included. There was a particularly high risk of lung cancer among never-smokers (RR = 9.03, 95% CI: 0.63–129.28) in the Six Cities Study even after adjusting for occupation, although this may have been a statistical anomaly resulting from the very small number of lung cancer deaths (8) among never-smokers.

Although our attempt to control for occupational exposure was constrained by the limitations in data quality, the findings nevertheless increase our confidence that the apparent increase in risk of general mortality—and in particular cardiopulmonary disease mortality associated with fine particle air pollution—was not the result of uncontrolled confounding by occupational exposures. In the ACS Study, even after the lung carcinogen index has been applied, the possibility of some residual confounding by occupation for mortality from lung cancer cannot be ruled out.

In both studies, occupational dirtiness rating exerted some effect modification. The air pollution effects tended to be stronger among subjects with high occupational dirtiness ratings, although the trends were not strictly monotonic. Education similarly was an effect modifier, and our attempts to disentangle the relative impacts of these two covariates did not produce a clear distinction. (It is important to remember that these two variables—education and occupational dirtiness—not only are correlated but also measure some of the same underlying social traits of the study subjects.)

Time-Dependent Covariates

In long-term cohort mortality studies, the values of important covariates may change over time, which leads to concomitant temporal changes in risk. Although all covariate values used in the ACS Study had been determined when the cohort was defined in 1982, longitudinal information on covariates was available for the Six Cities Study from the follow-up questionnaires administered at 3, 6, and 12 years after enrollment. Using Poisson regression,

Table 53. Relative Risks of Mortality from All Causes Associated with Selected Indices of Fine Particle Air Pollution[a] Based on the Multivariate Poisson Regression Model

Model	Fine Particle Index of Exposure	Relative Risk[b] (95% CI)
1	Exposure to $PM_{2.5}$ for each city remained fixed over the entire follow-up period	1.31 (1.13–1.52)
2	Exposure to $PM_{2.5}$ for each city was defined according to 13 calendar periods[c]	1.16 (1.02–1.32)
3	Exposure to $PM_{2.5}$ was assigned based on the city-specific mean exposure estimate for the earliest year available	1.19 (1.08–1.30)
4	Time-dependent estimate of $PM_{2.5}$ exposure received during the 2 years before death	1.16 (1.02–1.31)
5	Time-dependent estimate of $PM_{2.5}$ exposure received 3–5 years before death	1.14 (1.02–1.27)
6	Time-dependent estimate of $PM_{2.5}$ exposure received > 5 years before death	1.14 (1.05–1.23)

[a] Relative risks were calculated for a change in the pollutant of interest equal to the difference in mean concentrations between the most-polluted city and the least-polluted city; in the Six Cities Study, this difference for fine particles was 18.6 $\mu g/m^3$.

[b] Relative risks were adjusted for age, gender, body mass index, education, number of years smoked (at baseline), number of cigarettes smoked weekly, and occupational exposures.

[c] Exposures were defined according to 13 calendar periods: earlier than 1979, 1979, 1980, 1981, ..., 1989, and 1990 or later.

we found that incorporating information on changes over time in cigarette smoking and BMI had little effect on the association between fine particles and mortality. Allowing for the general downward trend in the average annual concentration of fine particles in the six cities, however, resulted in somewhat lower risk (RR = 1.16; 95% CI: 1.02–1.32) than we found with Poisson regression based on fixed-in-time long-term average fine particle levels (RR = 1.31; 95% CI: 1.12–1.52).

As discussed in the Time-Dependent Covariates section, the strong correlations that we observed between city-specific indices of fine particles did not allow us to discriminate among the risks of mortality for exposures received at various time intervals before death. For example, as shown in Table 53, the relative risk of mortality was roughly equivalent for exposures received within 2 years of death, 3 to 5 years before death, and more than 5 years before death. Multivariate models that included these indices simultaneously were highly collinear, so risk estimates were unstable (data not presented within this report). Further exploration of the variations in risk associated with fine particles is necessary before we can determine whether long-term or short-term exposure is most predictive of increased mortality. Such analyses will require detailed individual exposure information in which large changes in fine particle concentrations have occurred over an extended period of time.

An ideal analysis would include time-dependent exposure profiles for each individual in the study (Murdoch et al 1992). The construction of such profiles would require accurate information on study subject mobility linked to ambient fine particle monitoring data for each residence occupied during the period of interest, or even personal monitors. The development of residence histories and time-dependent exposures could be more informative for the ACS Study than for the Six Cities Study because the former exhibited greater variation in exposure patterns among participants and a larger number of persons who had moved.

Population Mobility

Population mobility is an important consideration in long-term follow-up studies, because cohort members may change residences, and hence change exposure, during the observation period. Mobility is particularly important in studies of environmental factors that affect population health, as the level of exposure may vary substantially with geographic location.

Population mobility is difficult to assess in the ACS Study, because subjects' residence changes were not generally monitored subsequent to 1982 enrollment. The Six Cities Study afforded a greater opportunity to assess mobility within the cohort. The Reanalysis Team constructed postrecruitment residence histories for cohort members using the follow-up interviews and the annual contacts with study participants.

Mobility within the Six Cities Study cohort was limited; only 18.5% of participants left their original city of residence during the follow-up period. The relative risk of mortality from all causes in the subcohort of nonmovers was similar to that in the entire cohort.

Movers were younger and better educated than nonmovers, and did not exhibit a significantly elevated relative risk of fine particle–associated mortality from all causes

(RR = 1.08, 95% CI: 0.67–1.76). However, relative risks declined with increasing educational attainment, decreasing from 1.56 (95% CI: 0.67–3.64) among movers with less than a high school education to 0.96 (95% CI: 0.40–2.30) among those with more than high school education. As the subgroup of movers was small, relative risks were estimated with less precision than in the much larger subgroup of nonmovers.

We also evaluated pre-enrollment mobility using the reported number of years subjects had lived in the original city of residence before they enrolled in the study. Including residency duration as a predictor of mortality from all causes did not appreciably alter the relative risk of mortality associated with exposure to fine particles.

Alternative Air Quality Data

In the Six Cities Study, the Original Investigators had monitored ambient air pollution levels throughout the study period using data from federal and state monitoring stations, as well data from their own monitors developed specifically for that study. In the ACS Study, 1980 data had been obtained for sulfate (151 cities) and fine particles (50 cities) from AIRS and from the EPA's IPMN. The Six Cities Study data have been subjected to several independent audits, including that by the Reanalysis Team. Our audit of the ACS air pollution data was more difficult because of the limited information about how the database was constructed.

In order to test the sensitivity of the relative risk estimates that had been obtained by the ACS Study Original Investigators, the Reanalysis Team developed several alternative indices of exposure to fine particle air pollutants. We examined all available AIRS data for the period 1980–1989, and constructed exposure indicators for 133 of the ACS Study cities from fine particle air pollution data for 1980–1981. With the AIRS data and additional data from the IPMN, we were able to assemble alternative sulfate data for 144 of the 151 cities in the ACS Study. These alternative sulfate data led to risk estimates similar to those obtained by the Original Investigators. However, correcting the sulfate data for a known artifact in the high-volume samplers used to generate the AIRS data reduced the sulfate concentrations by approximately 50%, and somewhat increased the multiplicative risk estimates for all-cause and cardiopulmonary disease mortality. These alternative sulfate data reduced the estimate of lung cancer mortality associated with sulfate concentration from 1.33 (95% CI: 1.10–1.61) using the original sulfate data to 1.18 (95% CI: 0.96–1.47) using the adjusted sulfate data.

Using data derived from the IPMN, we obtained fine particle measurements for 63 cities, rather than the 50 cities in the original ACS Study's fine particle cohort. These data led to estimates of risk slightly lower than those obtained by the Original Investigators for all-cause, cardiopulmonary disease, and lung cancer mortality.

Further analysis by the Reanalysis Team failed to reveal increased risks of mortality for inhalable particles (PM_{15}), the coarse particle fraction ($PM_{15-2.5}$), or total suspended particles in the approximately 60 cities for which such data were available from the IPMN. We noted no increased association between all-cause mortality and total suspended particles in the 154 cities for which total suspended particle data were available from AIRS.

Gaseous Copollutants

Air pollution is a complex mixture of not only fine particles and sulfate, but also gaseous copollutants including carbon monoxide, nitrogen dioxide, ozone, and sulfur dioxide. These gases, present in varying degrees in virtually all urban centers in the United States, are often highly correlated both spatially and temporally. They have been associated with cardiorespiratory morbidity and mortality in time-series studies, and it is possible that long-term exposure to these gases also contributes to the observed association between mortality and exposure to either fine particles or sulfate. Because of the strong interrelations among these copollutants, it is difficult to separate their effects. This is recognized as an area of high priority for future research (National Research Council 1998, 1999).

The Six Cities Study, with its small number of cities and high degree of correlation among the air pollutants monitored, did not permit a clear distinction among the effects of gaseous and fine particle pollutants. Indeed, estimates of the relative risk of mortality from all causes were similar for exposure to fine particles, sulfate, sulfur dioxide, and nitrogen dioxide. Of the gaseous copollutants in the Six Cities Study, only ozone did not display an association with mortality.

The ACS Study, which involved a much larger number of cities with more diverse ambient air pollution profiles, afforded a greater opportunity to evaluate the effects of the gaseous copollutants. The supplementary data assembled by the Reanalysis Team on sulfur dioxide, ozone, nitrogen dioxide, and carbon monoxide permitted us to roughly evaluate the impact of these gaseous pollutants on mortality. Although no positive associations were found in the Cox regression models between ozone, nitrogen dioxide, or carbon monoxide and mortality from all

causes, cardiopulmonary disease, or lung cancer, sulfur dioxide did demonstrate a significant association with all-cause and cardiopulmonary disease mortality. In the ACS Study, the association between sulfur dioxide and mortality persisted after we made adjustments for spatial autocorrelation (see below).

Sulfur Dioxide

We observed a stronger association between sulfur dioxide levels and mortality from all causes in the ACS Study than between either fine particles or sulfate and all-cause mortality. This difference in the strength of the association with mortality could result from the stability of the city-specific pollutant exposure estimates. Sulfate and fine particle mass data were obtained only every six days, whereas the gaseous pollution data were obtained hourly and averaged daily. Thus, city-specific average concentrations for gaseous pollutants comprised six times as many observations as the fine particle averages.

We therefore constructed a new exposure measure for sulfur dioxide, the gaseous pollutant most strongly associated with mortality. We used only those days in 1980 in each city for which there was also an available sulfate measurement. On the basis of this limited dataset, the relative risk of all-cause mortality associated with sulfur dioxide was 1.32 (1.24–1.40), a value similar to that based on all available observations (RR = 1.30; 95% CI: 1.23–1.38). Thus, the fact that sulfur dioxide was a stronger predictor of mortality than was sulfate does not appear to be due to the larger number of sulfur dioxide measurements.

We examined the association between mortality and exposure to sulfur dioxide for subjects who had not completed high school (RR = 1.28; 95% CI 1.15–1.43), subjects who had completed high school (RR = 1.50; 95% CI 1.35–1.67), and subjects who had had more than high school education (RR = 1.17; 95% CI 1.07–1.29). The sulfur dioxide effect on mortality risk was diminished for the best-educated subjects, a pattern we also observed with exposures to fine particles and sulfate. However, the sulfur dioxide effect, unlike the fine particle effect, was not the strongest for the least-educated subjects.

Acid Aerosols

Acid aerosols may mediate the association between fine particle air pollution and adverse health outcomes (Spengler et al 1990; Lippmann and Thurston 1996). We found an association with acid aerosols (RR = 1.12; 95% CI: 0.97– 1.30) in the Six Cities Study. However, we could not test this hypothesis for the ACS Study because no measurements of acid aerosols were available. In the two time-series studies reporting exposure to acid aerosols, no associations were found (Dockery et al 1992; Schwartz et al 1996). Lippmann and Thurston (1996) found that sulfate correlated well with acid aerosols; therefore, because sulfate-associated mortality risks were present in both studies under review here, an association between mortality and acid aerosols may exist for the ACS Study as it does for the Six Cities Study.

Ecologic Covariates

Gamble (1998) suggested that ecologic covariates other than the gaseous copollutants may have confounded the relation between fine particles and mortality in the ACS Study data. To address this concern, the Reanalysis Team considered 20 such ecologic covariates. (This analysis was necessarily restricted to the ACS Study; the inclusion of even five ecologic covariates other than air pollution in analyses of the Six Cities Study data would have resulted in a saturated risk model, in which the number of city-level covariates equals the number of cities.)

In order to avoid introducing ecologic covariates that may be artifactually associated with mortality, we identified a list of a priori covariates for which there was some plausible basis for suspecting an association with mortality. In this regard, we considered the main determinants of population health, including genetic, biological, environmental, occupational, social, and behavioral determinants, as well as health services. We obtained covariate values for the relevant metropolitan areas in the ACS Study from publicly available data sources such as the US Census Bureau and the US National Oceanic and Atmospheric Administration (NOAA).

When they are included in the Cox regression model, a number of these ecologic covariates appeared to be correlated with mortality. Specifically, population change, income, unemployment, education, income disparity, number of hospital beds, temperature, temperature variation, water hardness, sulfur dioxide, nitrogen dioxide, and ozone were significantly associated with mortality in the sulfate cohort ($P < 0.05$). However, because income disparity and nitrogen dioxide were negatively correlated with mortality, and water hardness was positively correlated, these ecologic associations require careful interpretation.

The diminished effect of sulfate on mortality risk estimates after we adjusted for either population change or exposure to sulfur dioxide warrants some discussion. The statistical effects of including ecologic covariates in the Cox proportional-hazards regression model are not well understood. Consequently, the Reanalysis Team investigated these effects by computer simulation under controlled conditions in which the true effect of the ecologic covariate is known (personal communication from Michal Abrahamowicz to the Ranalysis Team 2000). We found that including ecologic covariates in the Cox regres-

sion model reduced the estimated relative risk of the exposure of primary interest (in this case, fine particle air pollution). Unlike the case in linear ecologic regression, this effect was observed regardless of the strength of the association between the ecologic covariate and mortality, or the correlation between the ecologic covariate and the primary exposure. This downward bias in the relative risk of the primary exposure persisted even when the correlation between the covariate and exposure was negligible. This bias was small in many circumstances, but could be 20% or higher when the ecologic covariate was highly correlated with air pollution. The most important factor in determining the extent of this downward bias was the strength of the association between the covariate and mortality. Still, to obtain as accurate an estimate as possible of the sulfate-associated relative risk of mortality, our results suggest that it is better to adjust for relevant ecologic covariates than not.

Contextual Ecologic Effects

The Reanalysis Team made extensive efforts to identify ecologic covariates that would contribute to spatial variation in mortality rates among the cities in the ACS Study, and might confound the association between mortality and fine particle air pollution. Our multivariate Cox regression models identified a number of ecologic variables that altered the relative risk of mortality associated with fine particles or sulfate by more than 25%. Of the variables flagged as potential confounders in this way, population change, altitude, and sulfur dioxide were significant predictors of mortality in several of our multivariate Cox regression models. Although individual-level covariates, as had been used by the Original Investigators, are generally preferred over ecologic covariates in epidemiologic analysis, we used ecologic information on income because individual income was not recorded for the ACS Study. Variables such as population change and unemployment rate inherently represent community-level or contextual effects.

Certain variables can influence mortality at both the individual and the community level. For example, educational attainment was included as an individual covariate in our models, but also in our two-stage spatially filtered multivariate regression models for all-cause and cardiopulmonary disease mortality in the sulfate cohort.

Although this study focused on the association between fine particle air pollution and mortality, the contextual effects of education and other ecologic covariates on mortality warrant further exploration elsewhere. Useful guidance on the interpretation of the contextual education effect can be derived from recent work in the Netherlands (Schrijvers et al 1999). This study showed that inequalities in health, associated with inequalities in educational attainment, can be explained by a combination of behavioral and material factors. Behavioral factors include smoking, drinking, exercise, and dietary habits. Material factors include environmental quality (broadly defined to include housing, noise, and other pollution exposures) and factors such as the psychosocial stress of struggling to survive with the knowledge that others have much more material wealth. The relative risk of dying appears to take a fairly consistent ordering across educational strata and so do the material and behavioral risk factors. Behavioral and material factors together contribute to the health inequalities, but material factors contribute both directly (eg, through poor housing and environments) and indirectly (through modified behavior), which increases the importance of material factors.

This and other studies have suggested at least two possible explanations of the educational effect. One is differential exposure; less-educated persons appear less likely to avoid the risks of ambient air pollution (Hamilton 1995), and other factors that may influence health, than persons with higher education (Link and Phelan 1996). In a recent, ongoing study in Hamilton, Ontario, Jerrett (1999) found that enumeration areas (similar to the census block units in the United States) with high proportions of individuals with less than high school education were 2.5 times more likely to be within high pollution zones than areas populated more by educated individuals. This effect persisted even when researchers controlled for other potential predictors of exposure, including income and dwelling value.

The second possible explanation for the educational effect involves what can be called a healthy/unhealthy contagion phenomenon. Economic geographers use this concept to help explain the diffusion of technological innovation within specific regions (Miron 1984), and medical geographers use the term to characterize the movement of disease over space (Cliff and Haggett 1988). This concept might also be relevant for lifestyle behavior. Possibly, within better-educated communities, there is a propensity toward healthier living simply because people who live in close proximity are likely to imitate good behavioral traits in others around them. In other words, the interaction creates what could be a healthy community effect. The opposite, an unhealthy community effect, may possibly occur in less-educated communities. We would expect individuals in healthy communities to be better able to cope with a range of health risks, including air pollution, because of the relative absence of other risk factors. Likewise, we would expect to see air pollution exert a significantly greater effect in unhealthy communities.

A growing and impressive body of literature using multilevel models suggests that both personal characteristics and place of residence can help to explain health inequalities (see, for example, Duncan and Jones 1996). In our reanalysis of the ACS Study, we found some evidence that education was important in both individual and ecologic regression models. In the individual-level models, the pollution effect diminished or became insignificant when the model was stratified by education, and the ecologic models suggested that a community-wide educational effect may be at work. The individual model findings, combined with the importance of the ecologic education variable, lend more support to the possibility that exposure tends to vary with education.

This effect modification may well be due to the combination of spatial variation in intracity distributions of pollution, segregation of neighborhoods with low education, and the resulting inequitable exposure to pollution for persons of low education and places of high unemployment. In this case education becomes a marker for exposure misclassification, and although some of the effect modification by education may still result from contagious healthy or unhealthy living associated with community-wide education, the majority of the effect modification still results from exposure to pollution.

The association between mortality and air pollution found in the ACS Study may have been influenced mostly by persons of relatively low educational status who live in areas of high pollution. This interpretation would join a growing body of literature on environmental justice or equity (Jerrett 1997; Institute of Medicine 1998) that connects socioeconomic and racial status to disproportionate pollution exposures. Although much remains to be investigated regarding the health effects of such exposures, indications are that at least part of the socioeconomic and racial inequalities in health observed in the United States and other countries arise from the higher exposures of disadvantaged groups.

Some variables that we expected would have a significant effect on mortality were not strongly associated in our analyses. For example, although still controversial (Judge et al 1998), many studies (Kawachi and Kennedy 1997; Kaplan 1996; Wilkinson 1996) have suggested that income disparity (measured in our analyses by Gini coefficients) is an important determinant of health in the United States and that lower-income groups consistently display higher levels of age-standardized mortality. In the United States, income inequalities often translate into spatial inequalities (Massey and Denton 1993); because of residential segregation, disadvantaged individuals are subject to greater exposures from ambient fine particles, which may lead to misclassification bias.

These spatial inequalities may translate into exposure inequalities and a higher degree of exposure misclassification within the MSAs. Yet the Gini coefficient was not strongly associated with mortality in the ACS cohort. In the few models in which the Gini coefficient did achieve significance, it was negatively correlated. This unexpected finding might arise from the different locations for income disparity and high relative risk of mortality associated with air pollution. In the ACS Study, income disparity centered mostly in the Deep South around Mississippi, and the relative risks, as mentioned, were highest in the Lower Great Lakes region.

Spatial Autocorrelation

In the original analyses of both the Six Cities Study and the ACS Study, the possibility of spatial autocorrelation in both mortality rates and exposure to fine particle air pollution had not been considered. In addition, other covariates used in both the original analyses and the reanalysis may have been autocorrelated spatially. The presence of spatial autocorrelation affects the statistical power of the associations, with positive spatial autocorrelation in the residuals increasing the likelihood of a false-positive finding. In the ACS Study, which spanned the breadth of the continental United States, the existence of regional differences could lead to spatial autocorrelation, and we conducted additional analyses to take this into account.

The spatial analysis focused on city-specific mortality rates adjusted for all individual-level covariates, but not air pollution, using a risk model selected from a large number of alternative models fit to the ACS Study data. We then examined the association between these adjusted city-specific mortality rates and ecologic covariates, including fine particle levels and gaseous copollutant concentrations, using spatial regression methods. We detected significant positive spatial autocorrelation for most covariates, including the adjusted mortality rates, ambient air pollutant levels, and supplementary ecologic covariates such as education.

The analyses conducted by the Reanalysis Team to account for spatial autocorrelation in the ACS Study data are summarized in Figure 22. Both the standard Cox model that had been used by the Original Investigators and our two-stage regression Independent Observations Model assume statistical independence among all observations, and are included as points of reference.

We considered five additional two-stage random effects model specifications to account for spatial patterns in the data. The Independent Cities Model acknowledges the possibility that subjects living in the same city are more likely to have similar mortality rates than subjects living in

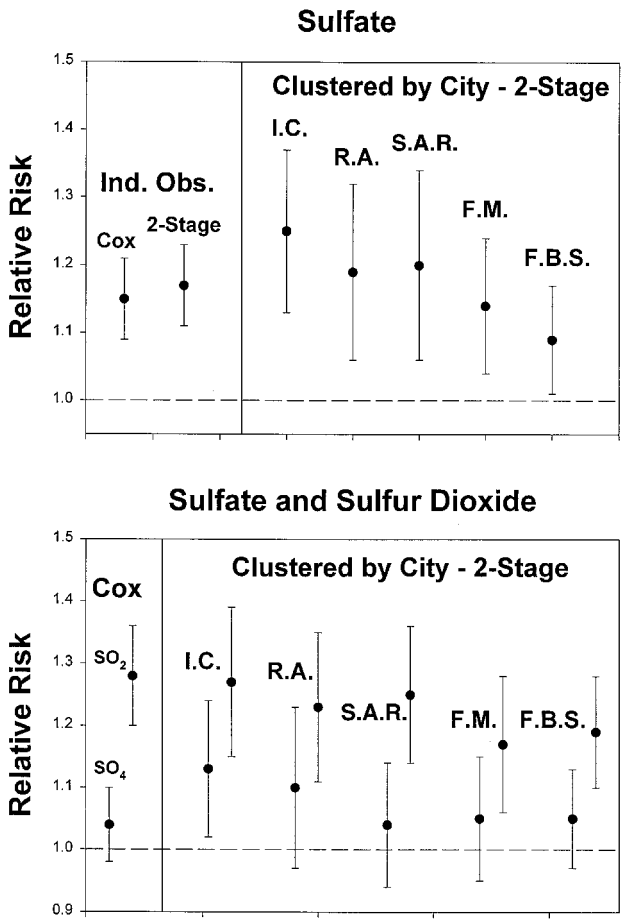

Figure 22. Effects of sulfate or sulfate and sulfur dioxide on relative risks of mortality from all causes using the ACS Study data. Top panel: Relative risks of mortality associated with an increase in sulfate concentrations of 19.9 μg/m^3 by type of statistical model for the dependency among subjects and cities. Bottom panel: Sulfate and sulfur dioxide were both included in each model and the relative risk of morality associated with each pollutant was calculated. In the left sections, the Independent Observations Models assume subjects are statistically independent (standard Cox model; two-stage regression assumes error variance equal to city-specific estimation error). In the right sections, all Clustered by City Models assume that the population's mortality is clustered within each city; risks were calculated using two-stage random effects regressions. In the Independent Cities Model (I.C.), data in different cities were assumed to be independent. In the Regional Adjustment Model (R.A.), mortality rates were adjusted for seven US regions. In the Simultaneous Autoregressive Model (S.A.R.), cities were assumed to be spatially dependent and analyses included an error structure based on nearest-neighboring cities. In the Filtered Mortality Only Model (F.M.), city-specific relative risks were spatially filtered before being linked with air quality data. In the Filtered Both Sides Model (F.B.S.), both relative risks and air quality data were spatially filtered. Error bars represent ±2 SE.

different cities. To look for evidence of clustering of mortality by city, we compared variation in the logarithms of the estimated city-specific relative risks obtained in the first stage of our two-stage regression approach to their standard errors. If the observed variation was similar to the estimation error, then cities could not be distinguished in terms of mortality rates. However, if there was additional variation in these estimates after we controlled for risk factors such as city-level air pollution measurements, then we modeled the variation in mortality rates among cities by including a random effect for each city with a common variance. This additional variation affects the estimates of the influence of city-level and ecologic covariates as well as their uncertainty.

In the first (Independent Cities) random effects model, we assumed that the mortality rates for cities are statistically independent. Under this model, the relative risk of all-cause mortality associated with exposure to airborne sulfate was increased over that in models assuming independence among subjects because both estimation error and variation in risk among cities were included in the weighting scheme for combining risk estimates across cities. This type of weighting scheme generates more uniform weights among cities than that used in the Independent Observations Model, in which the weights depend only on the estimation error. The degree of uniformity depends on the magnitude of the intercity variation in risk relative to the average within-city estimation error. Under the random effects model, if the variation among cities is comparatively large, then the city-specific weights will be similar and the larger cities are weighted less. If the mortality rates in these larger cities do not conform to a predictive model for sulfate-associated mortality, they will have less influence on the analysis, thus increasing the relative risk of mortality associated with exposure to sulfate.

Confidence intervals on the relative risk estimates for all-cause mortality obtained under the Independent Cities random effects model are much larger than those for the standard Cox model because the standard error of the logarithm of the sulfate relative risk was twice the size of that obtained from the Cox model. This indicates that we do not really have over half a million (the sample size) independent observations in the ACS Study with respect to determining air pollution effects; rather, the effective sample size in the ACS Study is approximately half the total sample size (ie, the number of subjects divided by the variance inflation factor of 2).

We then questioned the assumption that mortality risks were independent among cities. Even after we controlled for available risk factor information at the individual and ecologic level, there was evidence of residual spatial autocorrelation among the city-specific risk estimates. Thus there was some unidentified process that varied in space and resulted in broader regional clustering of mortality rates. We addressed the statistical form of this spatial dependence from four approaches. In the first (Regional Adjustment Model), we adjusted the city-specific relative

risks of seven regions of the United States by using indicator functions for each region in the two-stage random effects regression model. The sulfate-associated relative risk of mortality estimated under the Independent Cities Model (RR = 1.25) decreased by about 32% under this Regional Adjustment Model (RR = 1.19), suggesting that some of the association between sulfate and mortality was the result of broader regional associations in these variables. Statistical tests for spatial autocorrelation in the residuals from the Regional Adjustment Model provided little evidence of further spatial dependency in the data ($P > 0.10$).

We also attempted to remove spatial autocorrelation in the stochastic error structure of the model, as opposed to addressing spatial autocorrelation through the deterministic part of the model (as in the Regional Adjustment Model), using our simultaneous autoregressive modeling approach. In this case, we considered residuals to be spatially autocorrelated if they were a nearest-neighbor under the assumption that cities closer to one another have more similar mortality experiences than cities farther part. As with the Regional Adjustment Model, the sulfate-associated relative risk of mortality under the Nearest-Neighbor Model (RR = 1.20) was lower than that obtained from the Independent Cities Model (RR = 1.25). This modeling approach also eliminated much of the evidence of spatial autocorrelation in the residuals ($P > 0.10$).

Our third approach to the issue of spatial autocorrelation involved modeling spatial patterns in the data directly. Here, the relative risk of each city was compared with the risks for neighboring cities within a distance selected such that the residual spatial autocorrelation was minimized; we used all cities within a 600-km radius to determine an average risk for adjustment purposes. (This approach is similar to that for time-series mortality studies in which temporal trends in mortality rates are removed by using multiday moving-average filters.) This approach has the advantage over the Regional Adjustment Model that the data themselves effectively select the size of the regions and control for spatial autocorrelation at each city. To explore the impact of filtering only the mortality and not the pollution data (Filtered Mortality Only), we regressed sulfate concentrations on the spatially filtered city-specific mortality rates. The resulting sulfate-associated relative risk (RR = 1.14) was lower than it was under the other methods used. Thus, it appears that in the relation between sulfate and mortality there is subregional spatial structure that the Regional Adjustment Model did not remove.

The final approach (Filtered Both Sides) involved filtering not only the city-specific relative risks of mortality, but also the sulfate concentrations. Here, a 600-km radius was sufficient to remove any evidence of spatial autocorrelation in the sulfate data. With this approach, we compared mortality rates and sulfate levels after we removed broad spatial patterns in both variables, representing both sides of the regression equation relating mortality and air pollution. The estimated sulfate-associated relative risk was lower under the Filtered Both Sides Model (RR = 1.09) than under the Filtered Mortality Only Model (RR = 1.14), further suggesting that broad spatial patterns in sulfate concentrations account for some of the association with mortality.

These analyses provide strong evidence that mortality rates are clustered by city and that effects in neighboring cities are more similar than are those in distant cities. The spatial regression methods suggest that part of the relation between sulfate and mortality is probably due to some unobserved variable or group of confounding variables. In particular, we see that the sulfate-associated effect drops from a relative risk of 1.25 with the Independent Cities Model, to 1.19 with the Regional Adjustment Model, and to 1.09 with the Filtered Both Sides Model. Subtracting the results of the Regional Adjustment and Filtered Both Sides Models from the Independent Cities Model gives a possible range (RR of 0.6–0.16) over which the sulfate-associated effect results from spatial autocorrelation in the data. When we convert this to a percentage on the basis of the relative risk of 1.25 from the Independent Cities Model, it suggests that uncontrolled spatial autocorrelation accounts for 24% to 64% of the observed relation. Nonetheless, all our models continue to show an association between elevated risks of mortality and exposure to airborne sulfate.

Spatial Analysis of the Joint Effects of Sulfate and Sulfur Dioxide

Our standard Cox model analysis of ecologic covariates indicated that adjustment for sulfur dioxide, a gaseous copollutant, reduced the sulfate-associated relative risk of mortality. We examined the possible effects of spatial autocorrelation on this association, and the results are shown in the bottom panel of Figure 22. In all models considered, the relative risk of all-cause mortality associated with exposure to sulfur dioxide was greater than that for sulfate. Furthermore, when sulfur dioxide was included as a covariate, the sulfate-associated relative risk did not achieve formal statistical significance ($P < 0.05$) in any model except the Independent Cities Model. The relative risk associated with sulfur dioxide was statistically significant in all models examined. After we adjusted for sulfur dioxide levels, the variation in city-specific relative risks of mortality was much lower than after we adjusted for sul-

fate or fine particle concentrations. This suggests that sulfur dioxide accounted for much more of the variation in between-city mortality than sulfate. The addition of sulfur dioxide into the models with sulfate removed spatial autocorrelation in the residual mortality rates. However, the sulfur dioxide effect was also sensitive to the method of analysis; the relative risk for all-cause mortality based on the Independent Cities Model (RR = 1.27) exceeded that for the Filtered Both Sides Model (RR = 1.19). Thus, the effect observed at broad spatial levels (RR = 1.27 − 1.19 = 0.08) was less than that observed at local levels (RR = 1.19 – 1.00 = 0.19). In contrast, sulfate demonstrated a stronger effect at the broad spatial level (RR = 1.25 − 1.09 = 0.16) than at the local level (RR = 1.09 – 1.00 = 0.09). This may result from the degree of spatial autocorrelation in the air pollution data; using Moran *I*, the spatial autocorrelation for sulfur dioxide was 0.27 whereas that for sulfate was 0.39.

EXPOSURE MEASUREMENT ERROR

A potentially important source of exposure measurement error in both the Six Cities Study and the ACS Study is that data from fixed-site monitors rather than personal dosimeters necessarily were used to evaluate individual exposures to ambient air pollutants. For a number of cities in the ACS Study, data were available from more than one monitor, and those data were averaged to provide an indicator of exposure for all individuals in the city.

Air pollution exposure is an ecologic index that refers to cities and not individuals. As an ecologic variable, it has limitations associated with its use, including, among others, levels of air pollution before the studies began, variations within and between cities during follow-up, declining levels in ambient particles over time, and the changing chemical makeup of air pollution.

When fixed-site monitors are used and other internal sources of particles of comparable toxicity are excluded, measurement error should bias associations toward the null hypothesis of no effect and decrease statistical precision (Zidek et al 1996). If exposure measurement errors can be characterized, statistical methods exist to adjust risk estimates for those errors (Fung and Krewski 1999a,b). Empirical data on the exposure measurement error incurred from using fixed-site monitors to represent individual long-term average exposures are unavailable at present. Moreover, it will be difficult to generate such data because of the need to collect individual exposures over an extended period of time.

We attempted to gauge the potential impact of exposure measurement error on estimates of mortality risks associated with long-term exposure to fine particle air pollution by examining the variation among fixed-site monitors in the same metropolitan area. Intermonitor variation provided some information; specifically, for individuals with limited intracity mobility, the difference between the ambient pollutant concentration at the nearest monitor and the average of all city monitors provides a rough indication of the extent of exposure measurement error. However, the average of the multiple city monitors provides a better indication of long-term average exposure for highly mobile individuals within a city. This suggests that within-city intermonitor variation can roughly indicate the extent of exposure measurement error incurred when fixed-site ambient air pollution monitors are used instead of personal dosimeters. To the extent that there is a high degree of population mobility within a given city, this may actually overestimate the degree of exposure measurement error.

The Reanalysis Team employed within-city intermonitor variation as a rough indicator of exposure measurement error for fine particles to calculate, using the nonparametric simulation extrapolation method developed by Carroll and colleagues (1995), adjusted estimates of the relative risk of mortality based on a simplified Cox regression model that included fine particles and smoking. Because the degree of measurement error varied among cities, we prepared adjusted estimates for a range of possible degrees of measurement error.

The results indicate that this type of random exposure measurement error could lead to substantial underestimation of risk associated with long-term exposure to fine particle air pollution. In the Six Cities Study, for example, the estimated relative risk of 1.26 for all-cause mortality might be in the range of 1.30 (low measurement error) to 1.50 (high measurement error), if we could adjust for this source of error. However, because the true extent of exposure measurement error remains unknown, these adjusted risk estimates are only indicative, and they need to be interpreted with caution.

A more complete quantitative evaluation of the potential impact of exposure measurement error would require that additional sources of error be incorporated into the analysis. For instance, instrumentation error (both random and systematic) inherent in daily readings from the fixed-site monitors requires consideration. However, this source of error may be small in relation to the spatial variation in pollutant levels within a large metropolitan area. The complex interrelation between indoor and outdoor sources of exposure to fine particles also warrants consideration, as do time-activity patterns reflecting the time individuals spent outdoors.

It should be clear that, as in other studies of ambient air pollution, the estimates of increased mortality associated

with exposure to ambient air pollution exclude exposure to fine particles and other pollutants from indoor sources. The implicit assumption in such analyses is that total personal exposure can be partitioned into two components representing airborne fine particles from outdoor and indoor sources. Two questions that arise are whether this assumption is correct, and whether the effects observed in this and other studies were confounded by indoor air pollution.

With regard to the first question, fine and ultrafine airborne particles do penetrate indoors (Dockery and Spengler 1981a,b; Spengler et al 1981; Spengler and Sexton 1983), and individuals will be exposed to ambient fine particles regardless of their activity patterns. For example, Dockery and Spengler (1981a,b) estimated that the indoor concentration of fine particles of outdoor origin was about 70% of the outdoor value, although full air-conditioning could reduce this to about 30%.

Leech and colleagues (1996) estimated that North Americans spend almost 88.6% of their time indoors, 6.1% outdoors, and 5.3% in vehicles; thus one would expect indoor air to be important in these time-series studies. The level of exposure to fine particle air pollution from indoor sources varies by individual, depending on personal activity patterns. The main argument for partitioning the two components of fine particles is that particles generated indoors will be different from ambient air fine particles. Outdoor particles consist of coarse particles from dust, and finer fractions of sulfate and carbon particles generated mostly by internal combustion. Indoor air fine particles are generated by indoor sources and activities, such as cigarette smoke, radon, indoor combustion of fuels, molds, fungi, shedding of human skin, and personal grooming habits (Spengler and Sexton 1983; Wallace 1996).

If these levels are approximately constant across cities, then the differential effect on rates of mortality will be minimal and the indoor component should not confound the effects of the outdoor component. The fact that smoking status had little effect on the association between fine particles and mortality provides some support for this position. Neither the Six Cities Study nor the ACS Study included information on environmental tobacco smoke in the home and the workplace. However, because the effects of passive smoking on mortality are generally much smaller than the effects of active smoking, these effects could also be minor.

Although most cross-sectional studies have found very low correlations between personal exposures and indoor and outdoor levels of exposure to air pollution (Dockery and Spengler 1981b; Wallace 1996), a recent within-subject study suggested much higher correlations (Janssen et al 1998). The within-subject longitudinal component of variability answers the primary question of whether personal exposure, averaged across individuals, correlates with levels measured outdoors. This is very important in time-series studies, because the analysis focuses on day-to-day variations in air pollution, but it is not critical in an analysis that uses fixed values.

ACCURACY OF MORTALITY DATA

We have assumed, as did the Original Investigators, that the underlying causes of death were accurately reported and accurately coded. It has been found, however, that the accuracy of coding varies with cause of death (Alderson and Meade 1967; de Faire et al 1976; Engel et al 1980; Percy et al 1981). Cancer deaths usually are coded with more than 80% accuracy, but deaths from respiratory and cardiovascular diseases are often confused. In particular, when persons had these conditions concurrently and both contributed to death, there can be some uncertainty about which should be selected as the primary underlying cause. In other instances, there may be errors in selecting one underlying cause in a complex chain of health events (eg, cancer leading to pneumonia and then to respiratory failure). Largely for these reasons, the Original Investigators combined cardiovascular and respiratory diseases in their analyses. As part of our sensitivity analyses, we also conducted separate analyses for these causes of death. In the absence of differential errors between cities in reporting the underlying cause of death in either study, we would expect that such errors would dilute the true associations. Unfortunately, we have no data to confirm such an absence; obtaining such information would be a major undertaking and was outside of the scope of this project.

SELECTION BIAS

The results of the Six Cities Study and the ACS Study have influenced the development of national air pollution control policy in the United States. Therefore, it is important to consider the extent to which the studies' results are applicable to the general United States population.

There are two issues, one related to generalizability and the other to bias, regarding the representativeness of these findings. If study subjects were not representative of their entire communities, it could compromise the generalizability of the findings. For instance, if the study undersampled persons of low social class or some other socioeconomic or demographic indicator, then the findings strictly are applicable only to the proportion of the population that matches the profiles of the subjects who were included. In the ACS Study, for example, 94% of the population was white; thus the results may not apply to other racial groups. Still, it

can be argued that relevant biological processes are likely to be identical across racial groups, so that air pollution is likely to have similar effects on all segments of the population.

Another possibility is that study subjects were not representative of their target populations and differed in certain key characteristics from community to community. Statistical analyses could adjust for those differences for which individual data were available. However, these adjustments might not capture all important differences. Intangible and unmeasured factors, related perhaps to the sociodemographic profiles of the communities, could have a bearing on who was recruited into the studies. Statistical adjustments for such contextual ecologic effects could be attempted, although the completeness of such ecologic adjustments would be difficult to assure. A bias in estimating mortality effects from air pollution would occur only if the reasons for being included or excluded from the study differed city by city and if these reasons were also correlated with mortality rates.

In the Six Cities Study, potential subjects, selected from a sampling frame of addresses, had been included in the study if they were contacted successfully and agreed to participate. The methods used to select subjects from the sampling frames were based on sound statistical sampling procedures. If the structure or nature of the sampling frames (household voting lists or private census for commercial listings) differed from city to city, or if the fieldwork procedures varied in any way that could influence the likelihood of participation, then there could be differential subject enrollment in the different cities. Ferris and colleagues (1979) compared the age-gender distributions in the Six Cities Study to the 1970 US census, and found important deviations from the expected numbers of subjects enrolled in the various age strata as well as in some broad occupational groups. This suggests that subjects were not representative of the occupational and age distributions of the target populations in all cities. The authors concluded that the largest deviations occurred in Harriman.

Other than age, gender, and occupation, we have no information about the characteristics of the sampling areas and whether they were comparable to each other sociodemographically. Subjects had been enrolled over a 4-year period (1974 until 1979) and, unless those years saw secular changes in socioeconomic situations in these cities, we would not expect prolonged enrollment to affect the composition of the study population. Response rates (mean 77%) varied from a low of 73% in Watertown and Steubenville to a high of 81% in Portage and Harriman (Ferris et al 1979). This is a fairly narrow range and does not admit a great deal of latitude for serious discrepancies to occur between cities. Face-to-face interviews had been conducted with subjects. A bias could occur if study personnel had changed dramatically over this period and if different techniques had been used to elicit information from study subjects. Because most questions were not subject to interpretation, we would not expect this to be a major factor.

In the ACS Study, conversely, subjects had been enrolled by volunteers; standard statistical and epidemiological methods had not been used to select individuals

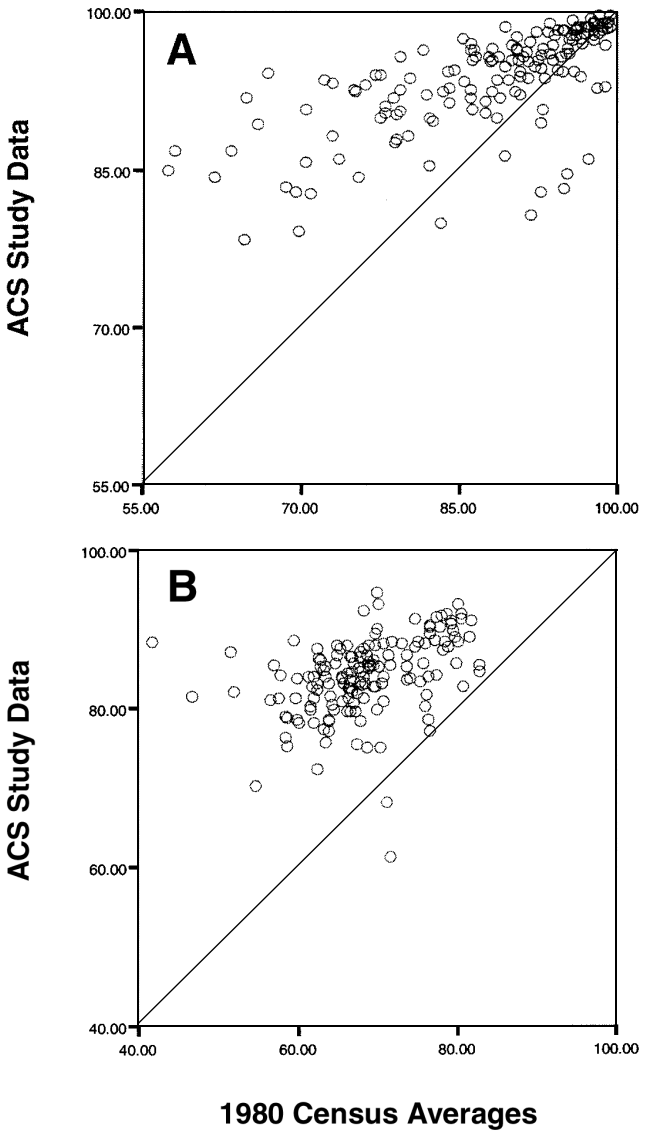

Figure 23. Comparison between ACS Study data and 1980 Census averages for race and educational attainment. A: Percentage of subjects (ACS Study) or residents (1980 US Census) in each MSA who defined themselves as being of white race. B: Percentage of subjects (ACS Study) or residents (1980 US Census) in each MSA who reported having completed high school.

from well-defined sampling frames. If different criteria and methods for enrollment had been used by the volunteers in different cities, it is conceivable that study participants in different cities varied substantially in their demographic, socioeconomic, and lifestyle characteristics. For example, volunteers in one city could have been, on average, more aggressive in persuading people to participate. As a result, the study participants in this city would comprise a broader psychological/social profile. In the city with less aggressive volunteers, the participants would be representative of the narrower, "easy to enroll", socially responsible segment. If these psychosocial characteristics were related to risk factors (such as education or occupation) and hence to mortality outcome, then intercity biases could ensue.

Because the ACS Study had no defined target population, response rates were also not defined, nor had any records been retained by ACS that could assist us in determining any city-specific biases. However, some evidence indicates that the ACS Study cohort is not representative. Figure 23a shows that the percentage of white persons enrolled in the ACS Study from each MSA is much higher than the census average for those same MSAs. Likewise, Figure 23b indicates that, according to 1988 Census Bureau data, the percentage of high school graduates was substantially higher among study participants than in the broader community of each MSA.

CONCLUSIONS

Time-series studies that focus on the effects of short-term exposure to fine particles, as well as cohort studies that address the effects of long-term exposures, have demonstrated significant associations between fine particle air pollution and mortality. This report has focused on the Six Cities Study and the ACS Study because they played a pivotal role in the establishment of the first NAAQS for fine particles in the United States. The importance of these two studies to regulatory policy development in the United States led to our independent audit and reanalysis, conducted for the Health Effects Institute.

In Part I of the reanalysis, we focused on validating the data that had been used by the Original Investigators in these two studies, and on replicating their numerical results using the same analytic methods. The data quality audit established the integrity of most of the data in both studies, with the exception of the air pollution monitoring data used in the ACS Study that had been obtained from third party sources and could not be validated. Although some analytic errors and discrepancies were noted in each study, these did not have a marked impact on the original risk estimates and did not materially affect the conclusions that had been reached by the Original Investigators.

In Part II, the Reanalysis Team conducted a detailed sensitivity analysis to assess the robustness of the original findings to alternative analytic approaches. We applied a wide range of alternative analytic approaches, including new methods of analysis such as random effects survival models and spatial filtering techniques. We also examined the effects of additional covariates taken from the original questionnaires that had not been included in the original published analyses, as well as the effects of 20 ecologic covariates that we developed from available databases and the general scientific literature.

The risk estimates reported by the Original Investigators were remarkably robust to alternative specifications of the underlying risk models, thereby strengthening confidence in the original findings. Specifically, the inclusion of additional individual-level covariates beyond those considered by the Original Investigators had little impact on the original risk estimates. Similar risk estimates also were obtained regardless of whether age or calendar year was used as the time axis.

The Reanalysis Team did find evidence of variation in risk among population subgroups; the most important was that the relative risk of mortality associated with fine particle air pollution decreased with increasing educational attainment. We observed this modifying effect of education in both studies. Although the interpretation of this finding is unclear, it is possible that educational attainment is a marker for socioeconomic status, which in turn may be correlated with exposure to fine particle air pollution.

In order to evaluate the possibility that the association between fine particles and mortality might result in part from occupational exposures, the Reanalysis Team developed and applied two new exposure indicators that measured occupational dirtiness and exposure to known lung carcinogens. These aggregate indicators of occupational exposure are particularly appropriate for respiratory conditions, malignant and otherwise, associated with inhalation of a range of substances that represent the exposures of most importance in occupational health risk assessment. Although cardiovascular disease has been associated with few occupational exposures, our overall dirtiness index may be indicative of exposure to workplace substances as yet unrecognized as increasing the risk of cardiovascular disease, and thus may afford some degree of control for occupational confounding in the case of cardiovascular disease as well as respiratory disease mortality.

Generally we found little evidence of uncontrolled occupational confounding of the association between fine

particle air pollution and mortality, but we could not rule out the possibility of residual confounding by occupation in the ACS Study with respect to the association between lung cancer mortality and exposure to sulfate. Our ability to adjust for occupational confounding in the ACS Study also was limited by the quality of the available occupational data.

In the Six Cities Study, allowing for changes in BMI and smoking during the follow-up period had little effect on the relative risk of mortality associated with fine particle air pollution. However, the relative risk of mortality from all causes decreased slightly when we accounted for the general decline in fine particle air pollution during the follow-up period, which suggests that the relative risk may be changing with time. The flexible exposure-response models applied by the Reanalysis Team also provided some evidence that the effects on mortality of both fine particles and sulfate were not constant over time.

Our analysis of residential mobility in the Six Cities Study indicated that only 18.5% of subjects moved from their original city of residence during the follow-up period. The risk estimates for this stable subcohort of nonmovers were similar to those for the entire cohort. Risk declined with increasing educational attainment in both the nonmover and mover subcohorts, even though the much smaller subcohort of younger, better-educated people that moved out of their original city of residence had not demonstrated an excess relative risk overall. These analyses could only be conducted with the Six Cities Study data, because temporal information on covariates was not available for the ACS Study.

The original air pollution monitoring data used in the ACS Study also were not available for a detailed audit; thus the Reanalysis Team considered a number of alternative indicators of exposure to air pollution in the sensitivity analyses of that study. Our measures of fine particles and sulfate correlated highly with the measures that had been used by the Original Investigators, and led to comparable relative risks of mortality from all causes, cardiopulmonary disease, and lung cancer, further increasing confidence in the original analysis. However, adjustment for a known artifact in the sulfate measurements reduced the indicators of sulfate exposure by about 50%, leading to an increase in the relative risk of all-cause and cardiopulmonary disease mortality associated with sulfate, but not in the relative risk of lung cancer mortality.

The inclusion of additional ecologic covariates in the ACS Study led to a number of new findings. Although adjustment for most ecologic covariates, in the Extended Model using Cox proportional-hazards regression, did not markedly affect the relative risks of mortality associated with fine particle air pollution, the inclusion of certain sociodemographic covariates (population change, in particular) reduced the relative risks for both fine particles and sulfate. Whereas the gaseous copollutants nitrogen dioxide, ozone, and carbon monoxide were not associated with mortality, sulfur dioxide was a significant predictor of mortality. Furthermore, adjusting for sulfur dioxide greatly diminished the effect of sulfate, and also somewhat reduced the association between fine particles and mortality. The roles of both sulfate and sulfur dioxide as predictors of mortality in the ACS Study support the notion that mortality may be related to more than one component of the complex mixture of urban air pollution in the United States. (Whereas sulfate levels reflect broader regional exposure conditions, sulfur dioxide levels are determined more by local point sources of air pollution.) The absence of a plausible toxicological mechanism by which sulfur dioxide could lead to increased mortality further suggests that it might be acting as a marker for other mortality-associated pollutants.

Because the original standard Cox model analyses of the ACS Study data had been predicated on the assumption that all observations are statistically independent, we conducted a number of analyses that allow for the spatial autocorrelation that was detected in the ACS Study data. These analyses employed two-stage random effects regression models that allowed for spatial clustering in mortality first at the city level, and then within seven broad airshed regions defined in the National Morbidity and Mortality Air Pollution Study (Samet et al 2000). Allowing for intracity or intraregional correlation in mortality resulted in slightly increased risk estimates that were subject to somewhat greater uncertainty than the original risk estimates.

We conducted additional spatial analyses after filtering out broader spatial patterns in mortality alone, or after filtering both the mortality and sulfate data, and obtained risk estimates comparable to or slightly lower than the risk estimates that had been reported by the Original Investigators. (Because of the limited number of cities for which measurements were available, spatial filtering could not be conducted for fine particle data.) The filtered risk estimates had wider confidence limits than the original risk estimates did, although the lower 95% confidence intervals for mortality from all causes and from cardiopulmonary disease were greater than the null value of unity.

Overall, these results, which allow for varying levels of spatial autocorrelation in the ACS Study data, support the association between fine particles and mortality that had been reported by the Original Investigators. However, the spatially adjusted risk estimates are subject to somewhat

greater uncertainty than the original risk estimates because of significant spatial autocorrelation in the ACS Study data.

The inclusion of ecologic covariates in our spatial adjustment models generally had somewhat less impact on the association between mortality and fine particles than did their inclusion in the original Cox regression model, which assumes no spatial autocorrelation in the data. In the Cox model, for example, the addition of population change, which demonstrated a high degree of spatial autocorrelation and a strong east-west gradient, notably reduced the relative risk of all-cause mortality associated with exposure to sulfate. In our spatial adjustment models, however, the sulfate-associated relative risk of all-cause mortality decreased only slightly when population change was included. In contrast, population change was a strong predictor of mortality in the standard Cox regression model, which assumes all individual observations are statistically independent.

Our spatial analyses also demonstrated a significant association between sulfur dioxide and mortality. Furthermore, this association appeared to be robust against adjustment for other ecologic covariates, including fine particles and sulfate, the covariates of primary interest here. In contrast, the inclusion of sulfur dioxide in our spatial regression methods resulted in a reduction in the mortality risk associated with exposure to both fine particles and sulfate. Nonetheless, even after adjustment for the effects of sulfur dioxide, both fine particles and sulfate demonstrated a positive association with mortality in our spatial regression methods.

Collectively, our reanalyses suggest that mortality may be associated with more than one component of the complex mix of ambient air pollutants in urban areas of the United States. Most of the individual pollutants that had been measured in the Six Cities Study demonstrated associations with mortality of similar magnitude, because the individual pollutants in these cities were highly correlated. Throughout the reanalysis of the ACS Study, both fine particles and sulfate demonstrated positive associations with mortality, as did sulfur dioxide.

Finally, it is important to bear in mind that the results of our reanalysis alone are insufficient to identify causal relations with mortality. Rather, we can conclude only that urban air pollution is associated with increased mortality in these two important epidemiologic investigations.

REFERENCES

Abbey DE, Colome SD, Mills PK, Burchette R, Beeson WL, Tian Y. 1993. Chronic disease associated with long-term concentrations of nitrogen dioxide. J Expo Anal Environ Epidemiol 3:181–202.

Abbey DE, Mills PK, Petersen FF, and Beeson WL. 1991. Long-term ambient concentrations of total suspended particulates and oxidants as related to incidence of chronic disease in California Seventh-day Adventists. Environ Health Perspect 94:43–50.

Abbey DE, Nishino N, McDonnell WF, Burchette RJ, Knutsen SF, Lawrence Beeson W, Yang JX. 1999. Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. Am J Respir Crit Care Med 159:373–82.

Abbey DE, Ostro BE, Petersen F, Burchette RJ. 1995. Chronic respiratory symptoms associated with estimated long-term ambient concentrations of fine particulates less than 2.5 microns in aerodynamic diameter ($PM_{2.5}$) and other air pollutants. J Expo Anal Environ Epidemiol 5:137–159.

Ahrens W, Merletti F. 1998. A standard tool for the analysis of occupational lung cancer in epidemiologic studies. Int J Occup Environ Health 4:236–240.

Alderson MR, Meade TW. 1967. Accuracy of diagnosis on death certificates compared with that in hospital records. Br J Prev Soc Med 21:22–29.

Amsel J, Waterbor JW, Oler J, Rosenwaike I, Marshall K. 1982. Relationship of site-specific cancer mortality rates to altitude. Carcinogenesis 5:461–465.

Bailey TC, Gatrell A. 1995. Interactive Spatial Data Analysis. Longman Scientific & Technical, Essex, England.

Boffetta P, Kogevinas M, Simonato L, Wilbourn J, Saracci R. 1995. Current perspectives on occupational cancer risks. Int J Occup Environ Health 1:315–325.

Brenner MH, Greeland S, Savitz DA. 1992a. The effects of nondifferential confounder misclassification in ecologic studies. Epidemiology 3:456–459.

Brenner MH, Savitz DA, Jockel KH, Greeland S. 1992b. Effects of nondifferential exposure misclassification in ecologic studies. Am J Epidemiol 135:85–95.

Breslow NE, Day NE. 1987. Statistical Methods in Cancer Research, Vol II, The Design and Analysis of Cohort Studies. Number 82. International Agency for Research on Cancer, Lyon, France.

Burnett R, Krewski D. 1994. Air pollution effects on hospital admission rates: A random effects modeling approach. Can J Stat 22:441–458.

Burnett RT, Brook JE, Philips O, Cakmak S, Raizenne M, Stieb D, Vincent R, Özkaynak H, Krewski D. 1998. The association between ambient concentrations of carbon monoxide

and daily mortality in Toronto, Canada. J Air Waste Manage Assoc 48:689–700.

Burnett RT, Dales R, Krewski D, Vincent R, Dann R, Brook JR. 1995. Associations between ambient particaulate sulfate and admission to Ontario hospitals for cardiac and respiratory diseases. Am J Epidemiol 142:15–22.

Carroll RJ, Ruppert D, Stefanski LA. 1995. Measurement Error in Nonlinear Models. Monographs on Statistics and Applied Probability 63. Chapman & Hall, London England.

Christiani DC, Wegman DH. 1995. Respiratory disorders. In: Occupational Health: Recognizing and Preventing Work-Related Disease (Levy BS, Wegman DH, eds) pp 427–454. Little, Brown, Toronto, Canada.

Clayton CA, Perritt RL, Pellizzari ED, Thomas KW, Whitmore RW, Wallace LA, Özkaynak H, Spengler JD. 1993. Particle total exposure assessment methodology (PTEAM) study: Distributions of aerosol and elemental concentrations in personal, indoor, and outdoor air samples in a Southern California community. J Expo Anal Environ Epidemiol 3:227–250.

Cleveland WS, Devlin SJ. 1988. Robust locally weighted regression and smoothing scatterplots. J Am Stat Assoc 74:829–836.

Cliff AD, Haggett P. 1988. Atlas of Disease Distributions: Analytical Approaches to Epidemiological Data. Basil Blackwell, Oxford, England.

Cook DG, Pocock SJ. 1983. Multiple regression in geographical mortality studies, with allowance for spatially correlated errors. Biometrics 39:361–371.

County and City Data Book. 1988. US Government Printing Office, Washington DC.

de Faire U, Friberg L, Lorich U, Lundman T. 1976. A validation of cause-of-death certification in 1,156 deaths. Acta Med Scand 20:223–228.

DerSimonian R, Laird N. 1986. Meta-analysis in clinical trials. Control Clin Trials 7:177–188.

Dockery DW, Pope CA, Xu X, Spengler JD, Ware JH, Fay ME, Ferris BG, Speizer FE. 1993. An association between air pollution and mortality in six US cities. N Engl J Med 329:1753–1759.

Dockery DW, Schwartz J, Spengler J. 1992. Air pollution and daily mortality: Associations with particulates and acid aerosols. Environ Res 59:362–373.

Dockery DW, Spengler JD. 1981a. Indoor-outdoor relationships of respirable sulfates and particles. Atmos Environ 15:335–343.

Dockery DW, Spengler JD. 1981b. Personal exposure to respirable particulates and sulfates. J Air Pollut Control Assoc 31:153–159.

Duncan C, Jones K. 1996. Health-related behaviour in context: A multi-level modeling approach. Soc Sci Med 42:817–830.

Easton DF, Peto J, Babiker AGA. 1991. Floating absolute risk: An alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. Stat Med 10:1025–1035.

Environmental Systems Research Institute. 1999. US Places (24000+). Environmental Systems Research Institute, Redland CA.

Engel LW, Strauchen JA, Chiazze L, Heid M. 1980. Accuracy of death certification in an autopsied population with specific attention to malignant neoplasms and vascular diseases. Am J Epidemiol 111:99–112.

Evans RG, Stoddart GL. 1990. Producing health, consuming health care. Soc Sci Med 31:1347–1363.

Feinleib M, Kannel WB, Tedeschi CG, Landau TK, Garrison RJ. 1979. The relation of antemortem characteristics to cardiovascular findings at necropsy—The Framingham Study. Atherosclerosis 34:145–157.

Ferris BG Jr, Speizer FE, Spengler JD, Dockery W, Bishop YMM, Wolfson M, Humble C. 1979. Effects of sulfur oxides and respirable particles on human health. Am Rev Respir Dis 120:767–779.

Fung K, Krewski D. 1999a. Evaluation of quasi-likelihood, regression calibration, and SIMEX methods in logistic regression when one of the predictors is subject to additive measurement error. J Epidemiol Biostat 4:65–74.

Fung KY, Krewski D. 1999b. On measurement error adjustment methods in Poisson regression. Environmetrics 10:213–224.

Gamble JF. 1998. $PM_{2.5}$ and mortality in long-term prospective cohort studies: Cause-effect or statistical associations? Environ Health Perspect 106:535–549.

Geolytics Software. 1999. CensusCD 1980. Geolytics Inc, East Brunswick NJ.

Gérin M, Siemiatycki J, Desy M, Krewski D. 1998. Associations between several sites of cancer and occupational

exposure to benzene, toluene, xylene and styrene: Results of a case-control study in Montreal. Am J Ind Med 34:144–156.

Gérin M, Siemiatycki J, Kemper H, Begin D. 1985. Obtaining occupational exposure histories in epidemiologic case-control studies. J Occup Med 27:420–426.

Getis A, Ord C. 1996. Local spatial statistics: An overview. In: Spatial Analysis: Modeling in a GIS Environment (Longley P, Batty M, eds). GeoInformation International, Cambridge MA.

Grambsch PM, Therneau TM. 1994. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 81:515–526.

Greenland S, Morgenstern H. 1989. Ecologic bias, confounding, and effect modification. Int J Epidemiol 18:269–274.

Greenland S, Robbins J. 1994. Invited commentary: Ecologic studies—Biases, misconceptions, and counterexamples. Am J Epidemiol 139:747–760.

Griffith DA, Doyle PG, Wheeler DC, Johnson DL. 1998. A tale of two swaths: Urban childhood blood-lead levels across Syracuse, New York. Ann Assoc Am Geogr 88:640–645.

Hamilton JT. 1995. Testing for environmental racism: Prejudice, profits, political power? J Policy Anal Manage 14:107–32.

Institute of Medicine, Committee on Environmental Justice. 1998. Toward Environmental Justice: Research, Education, and Health Policy Needs. National Academy Press, Washington DC

Janssen NA, Hoek G, Brunekreef B, Harssema H, Mensink I, Zuidhof A. 1998. Personal sampling of particles in adults: Relation among personal, indoor, and outdoor air concentrations. Am J Epidemiol 147:537–547.

Jerrett M. 1997. Environmental equity in Canada: An empirical investigation into the income distribution of pollution in Ontario. Environ Planning A29: 1777–1800.

Jerrett M. 1999. A GIS-spatial equity analysis of particulate air pollution in Hamilton. Presentation to the Canadian Geographers Association. Lethbridge, Alberta, Canada.

Judge K, Mulligan JA, Benzeval M. 1998. Income inequality and population health. Soc Sci Med 46:567–579.

Kaplan GA. 1996. People and places: Contrasting perspectives on the association between social class and health. Int J Health Serv 263:507–519.

Kaplan GA, Pamuk ER, Lynch JW, Cohen RD, Balfour JL. 1996. Inequality in income and mortality in the United States: Analysis of mortality and potential pathways Br Med J 7037:999–1003.

Kawachi I, Kennedy BP. 1997. The Relationship of income inequality to mortality: Does the choice of indicator matter? Soc Sci Med 45:1121–1127.

Koenig JQ, Larson TV, Norris G, Claiborn C, Finn D, Zweidinger R, Lewtas J, Schwartz J. 1998. A Fine Particulate Composition Variability and Exacerbation of Asthma. HEI Particulate Matter Research Inventory. Health Effects Institute, Cambridge MA.

Koenig JQ, Larson TV, Hanley QS, Rebolledo V, Dumler K, Checkoway H, Wang SZ, Lin D, Pierson WE. 1993. Pulmonary function changes in children associated with fine particulate matter. Environ Res 63:26–38.

Kotchmar DJ, McMullen T, Hasselblad V. 1987. Adequacy of a single monitoring site for defining mean outdoor concentrations of fine particles in a demarcated residential community. J Air Pollut Control Assoc 37:377–381.

Krewski D, Burnett RT, Goldberg MS, Siemiatycki J. 1998. Analytic Plan for the Re-analysis of the Harvard Six Cities Study and the American Cancer Society Study of Air Pollution and Mortality. Institute of Population Health, University of Ottawa, Ottawa, Canada.

Langbein LI, Lichtman AJ. 1978. Ecological Inference. Sage Publications, Beverley Hills CA.

Leaderer BP, Naeher L, Jankun T, Balenger K, Holford TR, Toth C, Sillivan J, Wolfson JM, Kountrakis P. 1999. Indoor, outdoor, and regional summer and winter concentrations of PM_{10}, $PM_{2.5}$, SO_4^{2-}, H^+, NH_4^+, NO_3^-, NH_3, and nitrous acid in homes with and without kerosene space heaters. Environ Health Perspect 107:223–231.

Leech JA, Wilby K, McMullen E, Laporte K. 1996. The Canadian Human Activity Pattern Survey: Report of methods and population surveyed. Chronic Dis Can 17:118–123.

Link BG, Phelan JC. 1996. Understanding sociodemographic differences in health—The role of fundamental social causes. Am J Public Health 4:471–473.

Lipfert FW, Malone RG, Daum ML, Mendell NR, Yang CC. 1988. A Statistical Study of the Macroepidemiology of Air Pollution and Total Mortality. BNL 52122 US-404. US Department of Energy, Brookhaven National Laboratory, Upton NY.

Lippmann M, Thurston GD. 1996. Sulfate concentrations as an indicator of ambient particulate matter air pollution for health risk evaluations. J Expo Anal Environ Epidemiol 6:123–146.

Massey DS, Denton NA. 1993. American apartheid: Segregation and the making of the underclass. Harvard University Press, Cambridge MA.

Miron J. 1984. Spatial autocorrelation in regression analysis: A beginner's guide. In: Spatial Statistics and Models (Gaile GL, Willmott CJ, eds) pp 201–222. D Reidel Publishing, Boston MA.

Moolgavkar SH, Luebeck EG. 1996. A critical review of the evidence on particulate air pollution and mortality. Epidemiology 7:420–428.

Murdoch DJ, Krewski D, Wargo J. 1992. Cancer risk assessment with intermittent exposure. Risk Anal 12:569–577.

National Research Council. 1998. Research Priorities for Airborne Particulate Matter I: Immediate Priorities and a Long-Term Research Portfolio. National Academy Press, Washington DC.

National Research Council. 1999a. Health Effects of Exposure to Radon, BEIR VI. National Academy Press, Washington DC.

National Research Council. 1999. Research Priorities for Airborne Particulate Matter II: Evaluating Research Progress and Updating the Portfolio. National Academy Press, Washington DC.

Odland J. 1988. Spatial Autocorrelation. Sage Publications, Beverly Hills CA.

Özkaynak H, Thurston G. 1987. Associations between 1980 US mortality rates and alternative measures of airborne particle concentration. Risk Anal 7:449–460.

Percy C, Stanek E, Gloeckler L. 1981. Accuracy of cancer death certificates and its effect on cancer mortality statistics. Am J Public Health 71:242–250.

Piantadosi S, Byar DP, Green SB. 1988. The ecological fallacy. Am J Epidemiol 127:893–904.

Pope CA, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE, Heath CW. 1995. Particulate air pollution as a predictor of mortality in a prospective study of US adults. Am J Respir Crit Care Med 151:669–674.

Preston DL, Lubin JH, Pierce DA, McConney ME. 1993. Epicure. Hirosoft International Corporation, Seattle WA.

Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, Schwarz J, Zanobetti A. 2000. The National Morbidity and Mortality Air Pollution Study, Part II: Morbidity and Mortality from Air Pollution in United States. Research Report 94. Health Effects Institute, Cambridge MA.

Sawada M. 1999. Global spatial autocorrelation indices–Moran's I, Geary's C and the general cross-product statistic. University of Ottawa website www.uottawa.ca/academic/arts/geographie/lpcweb/sections1/reports/moransi/moran.htm.

Schrijvers CT, Stronks K, van de Mheen HD, Mackenbach JP. 1999. Explaining educational differences in mortality: The role of behavioral and material factors. Am J Public Health 89:535–540.

Schwartz J, Dockery DW, Neas LM. 1996. Is daily mortality associated specifically with fine particles? J Air Waste Manage Assoc 46:927–939.

Schwartz J. 1991. Particulate air pollution and daily mortality: A synthesis. Public Health Rev 19:39–60.

Schwartz J. 1994a. What are people dying of on high air pollution days? Environ Res 64:26–35.

Schwartz S. 1994b. The fallacy of the ecological fallacy: The potential misuse of a concept and the consequences. Am J Public Health 5:819–824.

Shyrock HS. 1976. The Methods and Materials of Demography. Academic Press, New York NY.

Siemiatycki J, Nadon L, Lakhani R, Béégin, D, Géérin M. 1991. Exposure assessment. In: Risk Factors for Cancer in the Workplace (Siemiatycki J, ed) pp 45–114. CRC Press, Baton Rouge LA.

Spengler JD, Brauer M, Koutrakis P. 1990. Acid air and health. Environ Sci Technol 24:946–956.

Spengler JD, Dockery DW, Turner W, Wolfson M, Ferris BG. 1981. Long-term measurements of respirable sulfates and particles inside and outside homes. Atmos Environ 15:23–30.

Spengler JD, Sexton K. 1983. Indoor air pollution: A public health perspective. Science 221:9–17.

US Bureau of the Census. 1988. County and City Data Book. US Government Printing Office, Washington DC.

US Environmental Protection Agency. 1996. Review of the National Ambient Air Quality Standards for Particulate Matter: Policy Assessment of Scientific and Technical Information. OAQPS Staff Paper. EPA/452/R-96-013. Office of Air Quality Planning and Standards, Research Triangle Park NC.

Vedal S. 1997. Ambient particles and health: Lines that divide. J Air Waste Manag Assoc 5:551–581.

Wallace L. 1996. Indoor particles: A review. J Air Waste Manage Assoc 46:98–126.

Wilkinson RG. 1996. Unhealthy Societies: The Afflictions of Inequality. Routledge, London, England.

World Health Organization. 1975. *International Classification of Diseases, Ninth Revision*, Vol 1. World Health Organization, Geneva, Switzerland.

Zidek JV, Wong H, Le HD, Burnett R. 1996. Causality, measurement error and multicollinearity in epidemiology. Environmetrics 7:441–451.

PART II APPENDICES AVAILABLE ON REQUEST

The following appendices may be downloaded in PDF format from the Health Effects Institute Web site (*www.healtheffects.org*). Hard copies may be requested by contacting HEI at 955 Massachusetts Avenue, Cambridge MA 02139 (phone, 617-876-6700; fax, 617-876-6709; email, pubs@healtheffects.org). Please give the full title of the Special Report, the Part II title, and the titles of the appendices you wish to request.

A. Quality Assurance Audit of the Data
B. Occupational Exposures
C. Flexible Modeling of the Effects of Fine Particles and Sulfate on Mortality
D. Alternate Air Pollution Data in the ACS Study
E. Selection of Ecologic Covariates for the ACS Study
F. Definition of Metropolitan Areas in the ACS Study
G. Values of the Ecologic Covariates
H. Spatial Analyses
I. Random Effects Cox Models

Acknowledgments

Daniel Krewski, Richard T Burnett, Mark S Goldberg, Kristin Hoover, Jack Siemiatycki,
Michael Jarrett, Michal Abrahamowicz, Warren H White, and Others

This reanalysis of the Six Cities and ACS Studies of the association between particulate air pollution and mortality was a complex undertaking, involving a large number of scientists representing a range of disciplines. The Reanalysis Team itself comprised 31 individuals from 13 institutions in Canada and the United States; the new Centre for Population Health Risk Assessment in the Institute of Population Health at the University of Ottawa served as the focal point for the project.

All members of the Reanalysis Team made unique contributions to this multidisciplinary effort and are acknowledged elsewhere by name. Although all team members made outstanding contributions, Alette Willis deserves special recognition not only for her painstaking technical work in reconstructing the 1980 boundaries of the metropolitan areas included in the ACS Study and assembling additional socioeconomic, demographic, and environmental data for use by the Reanalysis Team, but also for serving as technical editor and coordinator for various report drafts. Kate Keating and Paula Carty also assisted with technical editing of report drafts, all of which were prepared under tight timelines. Additional contributions were made by Kimberly Zartolas and Ashley Wong, who spent the summer of 1999 at Harvard University coding the residence histories for all subjects included in the Six Cities Study, and by Sylvie Mauviel, who provided secretarial and administrative services to team members.

The Reanalysis Team gratefully acknowledges the cooperation of the Original Investigators in both the Six Cities Study and the ACS Study. Dr Douglas W Dockery, Dr Frank E Speizer, and Martha Fay answered many questions from the Reanalysis Team about the Six Cities Study, as did Drs C Arden Pope III, Michael J Thun, and Eugenia Calle about the ACS Study. Site visits with both sets of Original Investigators early in the Reanalysis Project were also of great value. We could not have asked for more patience and cooperation throughout the two-year period during which the reanalysis was conducted.

The Reanalysis Team met with the HEI Expert Panel four times during the course of the project to review interim results and to discuss methodologic approaches to specific issues. Discussions with the Expert Panel were extremely valuable and served to clarify a number of critical analytic issues. The Reanalysis Team also benefitted greatly from formal comments provided by the Special Panel of the HEI Health Review Committee on drafts of our reports submitted to HEI. Although the Advisory Board did not have the same opportunities to review the work in progress, their comments on major analytic issues were of great value to the Reanalysis Team.

The important role of HEI staff members also needs to be acknowledged. Dr Aaron J Cohen was the research project manager at HEI and skillfully facilitated interactions among the Reanalysis Team, Expert Panel, and Advisory Board. We were pleased to work with such an experienced and knowledgeable research manager. Daniel S Greenbaum, President of HEI, made the Reanalysis Project a personal priority and was actively engaged throughout the process. Dr JoAnn Ten Brinke was the review project manager and oversaw the production of the final report at HEI, coordinating the efforts of the HEI Review Panel and the editorial staff in the final stages of the project. This was a complex undertaking, given the intensive level of review to which the draft and revised reports were subjected as well as the length and scope of the final report. The editorial staff at HEI, led by Virgi Hepner for this project, did an outstanding job of editing and preparing the final manuscript for publication.

A project of this magnitude could not have been completed in a timely manner without excellent administrative support. Howard E Garsh at HEI and Gilles Morier at the University of Ottawa handled the detailed administrative arrangements, permitting the Reanalysis Team to focus on the task at hand, knowing that budgets and schedules were being carefully monitored. Mariella Peca and Hélène l'Abbée worked closely with Gilles on administrative and financial project matters at the University of Ottawa.

A number of individuals less directly involved with the Reanalysis Project also need to be acknowledged. Bob Bilgrad at the National Death Index provided guidance on procedures to be followed in securing the approvals needed to work with data on the vital status of individual subjects. Drs John Bachman and John Vandenberg directed questions from the Reanalysis Team about monitoring data collected by the US Environmental Protection Agency to

the right scientists at EPA. We specifically acknowledge Jacob G Summers, who provided the Reanalysis Team with data from the Aerometric Information Retrieval System, and Jose M Sune, who provided us with data from the Inhalable Particle Monitoring Network.

About the Primary Authors

Daniel Krewski is currently professor of medicine and professor of epidemiology and community medicine at the University of Ottawa, where he is involved in a number of activities in population health risk assessment within the new Insitute of Population Health. Dr Krewski has also served as adjunt research professor of statistics in the Department of Mathematics and Statistics at Carleton University since 1984. Prior to joining the Faculty of Medicine at the University of Ottawa in 1998, he was director of Risk Management in the Health Protection Branch of Health Canada. While with Health Canada, he also served as acting director of the Bureau of Chemical Hazards and as chief of the Biostatistics Division in the Environmental Health Directorate. Dr Krewski obtained his PhD in statistics from Carleton University and subsequently completed an MHA at the University of Ottawa. His professional interests include epidemiology, biostatistics, risk assessment, and risk management. Dr Krewski is a Fellow of the American Statistical Association and the Society for Risk Analysis. He currently serves as a member of the US National Academy of Sciences' Board on Environmental Studies and Toxicology and as a member of the Academy's Committee on Research Priorities for Airborne Particulate Matter.

Richard T Burnett received his PhD from Queen's University in 1982 in mathematical statistics. He is now a senior research scientist with the Environmental Health Directorate of Health Canada, where he has been working since 1983 on issues relating to the health effects of outdoor air pollution. Dr Burnett's work has focused on the use of administrative health and environmental information to determine the public health impacts of combustion-related pollution using nonlinear random-effects models, time series and spatial analysis techniques.

Mark S Goldberg is an associate professor with the National Institute of Scientific Research at the University of Quebec in Laval and is adjunct professor at McGill University in Montreal. Dr Goldberg obtained degrees in physics (BS 1975) and epidemiology and biostatistics (MS 1985, PhD 1992) from McGill. He is currently investigating the short-term effects of air pollution on mortality and identifying susceptible subgroups. His research interests include occupational risk factors for female breast cancer, and delivery of treatment and other health services for breast cancer.

Kristin Hoover started Hoover Consultants in 1989 after a request for quality assurance consulting on the Natural Resources Damage Assessment for the *Valdez* oil spill. During the past 11 years, she has worked on many audits in toxicology, analytic chemistry, regulatory review, and quality assurance, especially the potential health effects of particulate and gaseous air pollution. Prior to consulting, Ms Hoover was corporate product safety manager for Pennwalt Corporation and a senior toxicologist at Arco Chemical Company. She holds degrees in natural resources and environmental biology and is senior author of the book Managing Conduct and Data Quality of Toxicology Studies, which was the culmination of a conference sponsored jointly by government and industry. Ms Hoover was a founding board member of the Society of Quality Assurance, which now counts over a thousand members.

Jack Siemiatycki is professor of epidemiology at the University of Quebec with a concurrent appointment at McGill University. His 150 publications and reports cover many methodologic and substantive issues, including health survey methodology, juvenile diabetes epidemiology, health effects of asbestos, and occupational causes of cancer. He is recognezed for the development of novel methods of exposure assessment for investigation of environmental/occupational causes of cancer. Dr Siemiatycki is a Distinguished Scientist of the Medical Research Council of Canada.

Michael Jerrett is assistant professor in the School of Geography and Geology, the Health Studies Program, and the Environmental Health Program at McMaster University, Hamilton, Ontario, Canada. He earned his PhD in geography from the University of Toronto. He specializes in spatial analysis of environment–health relationships. Dr Jerrett has designed and analyzed local, provincial and national level health and environment databases in Canada, the United States, Mexico, and Vietnam.

Michal Abrahamowicz is associate professor of biostatistics in the Department of Epidemiology and Biostatistics at McGill University and medical scientist in the Division of Clinical Epidemiology at Montreal General Hospital. His research interests include survival analyses, nonparametric regression, epidemiology of cancer, cardiovascular and rheumatologic diseases, cohort studies, and randomized clinical trials.

Warren H White is a senior research associate in the Chemistry Department at Washington University in St Louis. Educated as a mathematician at the California Institute of Technology and University of Wisconsin, Dr White spent several years teaching and researching in the area that has since become known as *chaos theory*. He returned to California Institute of Technology to work on the 1972–73 Aerosol Characterization Experiment (ACHEX), an early multiinvestigator field study of particulate air pollution in Los Angeles. Dr White has since continued to study ambient aerosols, their effects on radiative transfer and visibility, and their relationship to particle and precursor gas emissions. Dr White is a member of the EPA's Clean Air Science Advisory Committee and serves on National Academy of Science panels assessing research priorities for airborne particulate matter and tropospheric ozone.

Abbreviations and Other Terms

MEASURES OF PARTICLES AND SULFATE

$PM_{2.5}$	particulate matter 2.5 μm or smaller in aerodynamic diameter
$PM_{2.5}$(DC)	**mean** fine particle fraction from dichotomous samplers
$PM_{2.5}$(DC MD)	**median** fine particle mass concentration from dichotomous samplers
$PM_{2.5}$(OI MD)	**median** fine particle concentration used by the Original Investigators
PM_{10}	particulate matter 10 μm or smaller in aerodynamic diameter
PM_{15}	particulate matter 15 μm or smaller in aerodynamic diameter
PM_{15}(DC)	**mean** inhalable particle fraction from dichotomous samplers
PM_{15}(SSI)	**mean** inhalable particle fraction from high-volume SSI samplers
$PM_{15-2.5}$	the coarse particle fraction of particulate matter [15-μm particles minus 2.5-μm particles]
$PM_{15-2.5}$(DC)	**mean** coarse particle fraction from dichotomous samplers
SO_4^{2-}	sulfate
$SO_4^{2-}_{(\text{cb-adj region})}$	sulfate data for 1980–1981 inclusive, with region-specific adjustment for artifactual sulfate
$SO_4^{2-}_{(\text{cb-adj season})}$	sulfate data for 1980–1981 inclusive, with season-specific adjustment for artifactual sulfate
$SO_4^{2-}_{(\text{cb-adj US})}$	sulfate data for 1980–1981 inclusive, with US-specific adjustment for artifactual sulfate
$SO_4^{2-}_{(\text{cb-unadj})}$	sulfate data for 1980–1981 inclusive, unadjusted for artifactual sulfate
SO_4^{2-}(DC)	sulfate data from PM_{15}(DC)
SO_4^{2-}(OI)	sulfate data used by the Original Investigators
TSP	total suspended particles
TSP(IPMN)	**mean** TSP mass concentrations based on IPMN data

OTHER TERMS

ACS Study	the American Cancer Society Study
AIRS	Aerometric Information Retrieval System
ARRCCM	*American Review of Respiratory and Critical Care Medicine*
BMI	body mass index
$CaCO_3$	calcium carbonate
CAPITA	Center for Air Pollution Impact and Trend Analysis
CASAC	Clean Air Science Advisory Committee
CI	confidence interval
CO	carbon monoxide
CPS-II	American Cancer Society's Cancer Prevention Study II
DC	measurement from a dichotomous sampler
df	degrees of freedom
EPA	US Environmental Protection Agency
FP+CP	fine particles + coarse particles
FVC	forced vital capacity
H^+	aerosol acidity
HSPH	Harvard School of Public Health
IARC	International Agency for Research on Cancer
ICD-9	*International Classification of Diseases, Ninth Revision*
IP	inhalable particles
IPMN	Inhalable Particle Monitoring Network
JAWMA	*Journal of the Air and Waste Management Association*
MA	metropolitan area
MD	median
MSA	metropolitan statistical area
NAAQS	National Ambient Air Quality Standard
NAD	National Aerometric Database
NDI	National Death Index
NEJM	*New England Journal of Medicine*
NO_2	nitrogen dioxide

NOAA	US National Oceanic and Atmospheric Administration
O_3	ozone
OSI	Office of Scientific Integrity
r	bivariate correlation coefficient
range	the difference in mean concentrations between the most-polluted city and the least-polluted city
RR	relative risk
SAS	Statistical Application Software
SID	subject identification number
Six Cities Study	the Harvard Six Cities Study
SO_2	sulfur dioxide
SSI	high-volume sampler with size-selective inlet
SSN	Social Security Number

RELATED HEI PUBLICATIONS

PARTICULATE MATTER

Report Number	Title	Principal Investigator	Publication Date*
Research Reports			
75	Ozone Exposure and Daily Mortality in Mexico City: A Time-Series Analysis	DP Loomis	1996
81	Methods Development for Epidemiologic Investigations of the Health Effects of Prolonged Ozone Exposure	IB Tager	1998
82	Acute Effects of Ambient Ozone on Asthmatic, Wheezy, and Healthy Children	JM Peters	1998
83	Daily Changes in Oxygen Saturation and Pulse Rate Associated with Particulate Air Pollution and Barometric Pressure	DW Dockery	1999
86	Statistical Methods for Epidemiologic Studies of the Health Effects of Air Pollution	W Navidi	1999
91	Mechanisms of Morbidity and Mortality from Exposure to Ambient Air Particles	JJ Godleski	2000
93	Effects of Concentrated Ambient Particles in Rats and Hamsters: An Exploratory Study	T Gordon	2000
94	The National Morbidity, Mortality, and Air Pollution Study *Part I.* Methods and Methodologic Issues *Part II.* Morbidity, Mortality, and Air Pollution in the United States	JM Samet	2000
95	Association of Particulate Matter Components with Daily Mortality and Morbidity in Urban Populations	M Lippmann	2000
96	Acute Pulmonary Effects of Ultrafine Particles in Rats and Mice	G Oberdörster	2000
…	Daily Mortality and Fine and Ultrafine Particles in Erfurt, Germany *Part I.* Role of Particle Number and Particle Mass (in Press)	H-E Wichmann	2000
…	Identifying Subgroups of the General Population that May Be Susceptible to Short-Term Increases in Particulate Air Pollution: A Time-Series Study in Montreal, Quebec (in Press)	MS Goldberg	2000
…	A Case-Crossover Analysis of Fine Particulate Matter Air Pollution and Out-of-Hospital Sudden Cardiac Arrest (in Press)	H Checkoway	2000
Special Reports			
	Diesel Exhaust: A Critical Analysis of Emissions, Exposure, and Health Effects		1995
	Particulate Air Pollution and Daily Mortality: The Phase I Report of the Particle Epidemiology Evaluation Project		
	Phase I.A. Replication and Validation of Selected Studies		1995
	Phase I.B. Analyses of the Effects of Weather and Multiple Air Pollutants		1997
HEI Communications			
5	Formation and Characterization of Particles: Report of the 1996 HEI Workshop		1997
8	The Health Effects of Fine Particles: Key Questions and the 2003 Review		1999
HEI Program Summaries			
	Research on Particulate Matter		1999

** Reports published since 1990.*
Copies of these reports can be obtained by writing or calling the Health Effects Institute, 955 Massachusetts Avenue, Cambridge MA 02139. Phone (617) 876-6700 FAX (617) 876-6709 E-mail pubs@healtheffects.org www.healtheffects.org

HEALTH
EFFECTS
INSTITUTE

955 Massachusetts Avenue
Cambridge MA 02139 USA
+1-617-876-6700

SPECIAL REPORT

July 2000